LEVEL II

12

VSC-TR-81-17

USE OF BACK AZIMUTH MEASUREMENTS IN SEISMIC EVENT LOCATION WITH REGIONAL DATA

D.W. Rivers, J.A. Burnetti and A.C. Chang
Seismic Data Analysis Center
Teledyne Geotech
314 Montgomery Street
Alexandria Virginia 22314

16 OCT 1981

Monitored By:
VELA Seismological Center
312 Montgomery Street
Alexandria, VA 22314

Sponsored by

The Defense Advanced Research Projects Agency (DARPA)

DARPA Order No. 2551

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: VSC-TR-81-17	2. GOVT ACCESSION NO.: AD-A106 [illegible]
	3. RECIPIENT'S CATALOG NUMBER: [illegible]
4. TITLE *(and Subtitle)*: USE OF BACK AZIMUTH MEASUREMENTS IN SEISMIC EVENT LOCATION WITH REGIONAL DATA.	5. TYPE OF REPORT & PERIOD COVERED: Technical rept.
	6. PERFORMING ORG. REPORT NUMBER: SDAC-TR-81-3
7. AUTHOR(s): D. W. Rivers, A. C. Chang, J. A. Burnetti	8. CONTRACT OR GRANT NUMBER(s): F08606-79-C-0007, DARPA Order 2551
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Teledyne Geotech, 314 Montgomery Street, Alexandria, Virginia 22314	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: VT/0709/B/PMP
11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Advanced Research Projects Agency, 1400 Wilson Boulevard, Arlington, Virginia 22209	12. REPORT DATE: 9 January 1981 13. NUMBER OF PAGES: 30
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: VELA Seismological Center, 312 Montgomery Street, Alexandria, Virginia 22314	15. SECURITY CLASS. *(of this report)*: Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

seismology
location
3-component processor
L_g
back azimuth
confidence region

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Conventional methods for seismic event location, which are based on measurements of signal arrival times, may be inadequate for events detected at only regional distances, because there may be an insufficient number of observations. This shortcoming can be alleviated by supplementing the measurements of P-wave arrival times with measurements of the event's back azimuth made by applying Smart's (1978) three-component processor to the regional phase L_g. Back azimuth measurements can also be used as an analyst's tool to calculate an approximate epicenter for use in the association of mixed events with the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

AD-A108258

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

appropriate signal arrivals or for use as the initial value in an iterative location scheme. An algorithm is herein presented which performs event location by treating arrival time measurements and back azimuth measurements in an analogous fashion. Tests of this algorithm which were run in the depth-restrained mode for two events show that it can be effective for sparse data sets but also that it can lead to large errors if there is a poor geometrical distribution of detecting stations. Computation of the confidence region about the epicenter demonstrates an insufficiency in the theory for calculating F-statistic confidence ellipses for azimuths and travel-time data if the measurement variances are not known *a priori*. The theory is satisfactory for the standard chi-squared ellipses.

USE OF BACK AZIMUTH MEASUREMENTS IN SEISMIC EVENT LOCATION WITH REGIONAL DATA

SEISMIC DATA ANALYSIS CENTER REPORT NO.: SDAC-TR-81-3

AFTAC Project Authorization No.:	VELA /0709/B/PMP
Project Title:	Seismic Data Analysis Center
ARPA Order No.:	2551
Name of Contractor:	TELEDYNE GEOTECH
Contract No.:	F08606-79-C-0007
Date of Contract:	1 October 1980
Amount of Contract:	$1,493,393
Contract Expiration Date:	30 September 1981
Project Manager:	Robert R. Blandford (703) 836-3882

P. O. Box 334, Alexandria, Virginia 22314

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

ABSTRACT

Conventional methods for seismic event location, which are based on measurements of signal arrival times, may be inadequate for events detected at only regional distances, because there may be an insufficient number of observations. This shortcoming can be alleviated by supplementing the measurements of P-wave arrival times with measurements of the event's back azimuth made by applying Smart's (1978) three-component processor to the regional phase L_g. Back azimuth measurements can also be used as an analyst's tool to calculate an approximate epicenter for use in the association of mixed events with the appropriate signal arrivals or for use as the initial value in an iterative location scheme. An algorithm is herein presented which performs event location by treating arrival time measurements and back azimuth measurements in an analogous fashion. Tests of this algorithm which were run in the depth-restrained mode for two events show that it can be effective for sparse data sets but also that it can lead to large errors if there is a poor geometrical distribution of detecting stations. Computation of the confidence region about the epicenter demonstrates an insufficiency in the theory for calculating F-statistic confidence ellipses for azimuths and travel-time data if the measurement variances are not known *a priori*. The theory is satisfactory for the standard chi-squared ellipses.

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

Early work in the field of seismic monitoring of nuclear explosions assumed both the existence and the utility of seismograms which were recorded at relatively short distance from the explosion sites. A formal distinction between these close-in measurements and those recorded farther away was made at the 1958 "Conference of Experts" in Geneva, which defined three distance zones for seismic observations. Zone 1 consisted of the distance range $0° \leq \Delta \leq 10°$; Zone 2, the range $10° \leq \Delta \leq 25°$; and Zone 3, $25° \leq \Delta \leq 100°$. Although there was some doubt about the value (for the purposes of discrimination and yield estimation) of signals recorded in Zone 2, since such signals would pass mainly through the heterogeneous upper mantle, it was generally agreed that signals recorded in Zone 1 would be valuable, particularly for the analysis of small events. Failure to agree upon a scheme for conducting seismic monitoring within the potential boundaries of the country being monitored, however, coupled with the inherent difficulty in interpreting those signals whose propagation is confined to the crust, eventually led to the neglect of studies utilizing Zone 1 data in favor of other work which concentrated on the analysis of data recorded in Zone 3. Although this concentration on teleseismic data has persisted for many years, recent CTBT negotiations have once again raised the possibility of installing seismic monitoring stations inside the borders of the U.S.S.R., and attention is once again being paid to the utilization of data recorded in Zones 1 and 2. A number of recent studies have addressed the question of how to use regional data to perform seismic discrimination; in this report we shall examine a technique for using regional data to perform event location.

Location with regional data presents many problems beyond those normally encountered when teleseismic data are used. Rather than measuring a single P-wave arrival (and later distinct phases such as S, PP, PKP, etc.), at regional distances an observer will note the arrival of such crustal phases as P_n, P_g, $\overline{P}$, S_n, etc., which propagate with velocities which may vary strongly even along propagation paths close to each other. Typically, only a few stations are situated at regional distances from a given event (this would especially be true in the case of U.S.-operated monitoring stations within the U.S.S.R.), and these stations may all be located within a small range of azimuths. Chang and Racine (1979) and Chang et al. (1980) have presented methods for overcoming the variability of the travel-time relations by means of

a detailed analysis of the propagation paths and by bootstrap location algorithms; we herein describe a method which attempts to compensate for the paucity of arrival time measurements by supplementing those data with back azimuth measurements, which can be made at regional distances by applying a surface-wave processor (Smart, 1978) to three-component observations of the phase L_g.

USE OF BACK AZIMUTH MEASUREMENTS IN EVENT LOCATION

The traditional determination of hypocenters and origin times of seismic events involves finding the best values of the coordinates (latitude, longitude, depth, origin time) such that, at a network of detecting stations the signal arrival time differences $\delta t = t_{observed} - t_{calculated}$ are minimized in a least-squares sense. In order to carry out the location, it is necessary to have measured arrival times at four or more stations, in order that each of the four coordinates of the hypocenter can be determined uniquely. An accurate location, i.e., one for which the "error ellipsoid" formed by the confidence limits surrounding the calculated hypocenter is small, would obviously require several more than four measurements of arrival time. In practice, this requirement may be difficult to meet, and it may be advantageous to supplement the arrival time data with measurements of the azimuths of the incoming signal or can be done at seismic arrays. As has already been mentioned, one instance in which there may be an insufficient number of measured arrival times is the case of a small event which cannot be detected at teleseismic distances. Except in those parts of the world which are densely populated with seismic stations, such events are unlikely to be detected by enough stations at regional distances to permit them to be located accurately. Measuring back azimuths as well as arrival times would effectively double the number of observations, and the additional information might be sufficient to resolve the ambiguities introduced by the inadequate number of arrival times. Even if there do exist enough regional and teleseismic arrival time measurements to permit an event to be located, back azimuth measurements may still be helpful in calculating that location. An important example of the need for auxiliary information, even when enough arrival times were measured, is the rather common instance of mixed arrivals from different events. In such a case one may be hard pressed to associate each signal with the proper event, and the resulting locations will be erroneous if the wrong choice of event associations is made. Knowledge of the back azimuths of some of the signals, however, will probably resolve any ambiguity in their association with the proper event; the arrival times and azimuths of these signals may be used to compute provisional epicenters and origin times for the events, and the predicted arrival times for signals emitted at these provisional epicenters may then be used to resolve the ambiguities in the association of the remaining signals (for which no

back azimuth data are available). Back azimuth measurements from single stations should thus serve as a valuable analyst's tool (as they do at arrays) for computing trial locations to be used as initial guesses in an iterative scheme.

In order to profit from the advantages offered by the inclusion of back azimuth data in the calculation of event locations, particularly at regional distances, we have rewritten the SDAC program LOC to process back azimuth measurements in the same manner as it conventionally handles measurements of arrival time. This report is intended to document this change and to present the result of some simple tests of location using back azimuths. We shall first describe the geometrical foundation of the new algorithm, and then we shall describe those changes in the program code of which the user should be aware. Finally, we shall examine in detail two case histories which use regional data and the new technique for event location.

THE LOCATION ALGORITHM

As in the traditional approach to event location which uses arrival time data only, the revised version of the program LOC assumes an initial trial value of the hypocenter and origin time, and it uses them to calculate the predicted signal arrival time and back azimuth at each station in the network of detecting stations. The conventional technique computes the difference δt_i between the observed and predicted arrival times at each station, and then it expands these residuals in a Taylor series

$$\delta t_i = \left(\frac{\partial t}{\partial T}\right)_i dT + \left(\frac{\partial t}{\partial x}\right)_i dx + \left(\frac{\partial t}{\partial y}\right)_i dy + \left(\frac{\partial t}{\partial h}\right)_i dh. \qquad (1)$$

If the data base consists of n measurements of arrival times, there will be n equations of the form (1), and the overdetermined linear system of these n equations may be solved in a least-squares sense by means of matrix inversion to yield the corrections dT, dx, dy, and dh which are to be added to the assumed values of the origin time, east, north, and depth coordinates. The partial derivatives in equation (1) may be evaluated by setting $(\partial t/\partial T)$ equal to unity and computing $(\partial t/\partial \Delta)_i$ and $(\partial t/\partial h)_i$ from tables of travel times. If the data base also includes m measurements of back azimuth, the revised algorithm supplements the system (1) with m equations of the form

$$\delta \zeta_i = \left(\frac{\partial \zeta}{\partial x}\right)_i dx + \left(\frac{\partial \zeta}{\partial y}\right)_i dy \qquad (2)$$

where $\delta \zeta_i$ is the difference between the observed and predicted values of the back azimuth at the i^{th} station. We note that back azimuth measurements provide no information about the origin time or the depth of the event. The expanded system of m+n equations may now be inverted as before to yield dT, dx, dy, and dh.

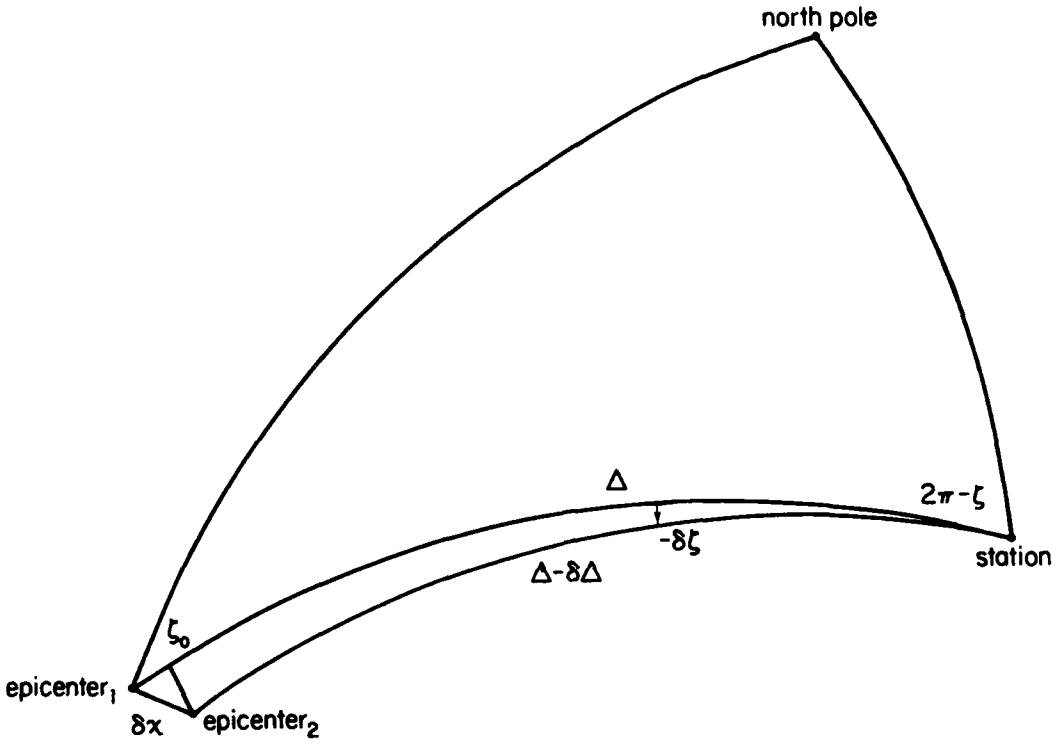
Evaluation of the partial derivatives in equation (2) is made straightforward by constructing the spherical triangles having one vertex at the detecting station and the other vertices at the trial epicenters both before and after displacement by δx and by δy (Figures 1a and 1b). Denoting the angular displacements by $\delta x = \delta x \cdot 0.0090$ deg/km and $\delta y = \delta y \cdot 0.0090$ deg/km and the forward azimuth by ζ_o we find

$$-(\delta \Delta \cdot 111.1 \text{ km/deg})/\delta x = \sin \zeta_o \qquad (3)$$

$$\sin(-\delta \zeta)/\sin \delta x = \sin(\pi/2 - \zeta_o)/\sin(\Delta - \delta \Delta) \approx \cos \zeta_o / \sin \Delta \qquad (4)$$

$$-(\delta \Delta \cdot 111.1 \text{ km/deg})/\delta y = \cos \zeta_o \qquad (5)$$

$$\sin(+\delta \zeta)/\sin \delta y = \sin \zeta_o / \sin(\Delta - \delta \Delta) \approx \sin \zeta_o / \sin \Delta. \qquad (6)$$



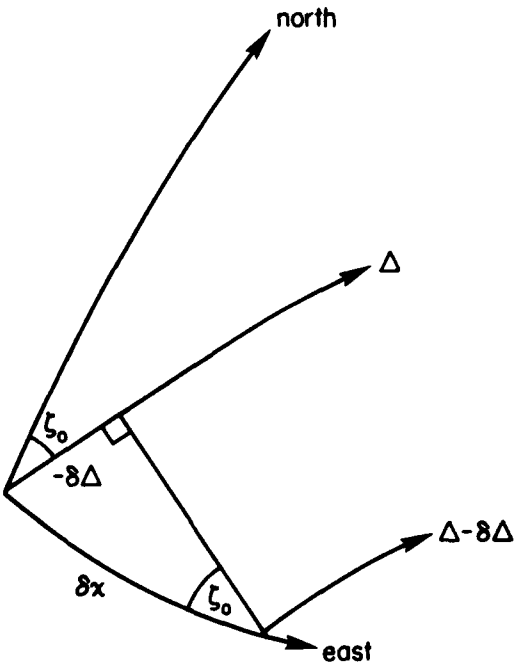


Figure 1a Effect on distance and back azimuth of displacing the trial epicenter. Angular displacement δ x eastward.

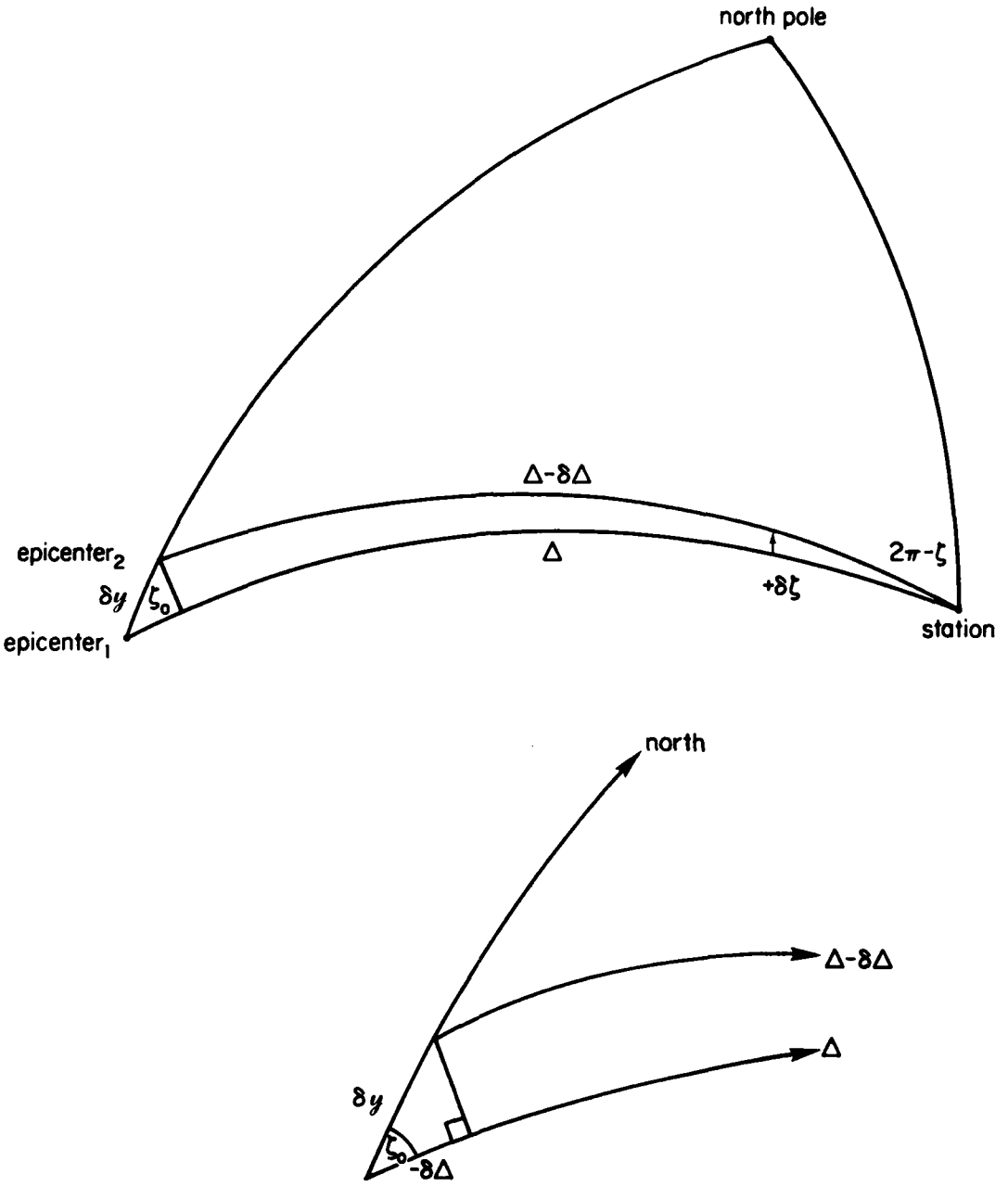


Figure 1b Effect on distance and back azimuth of displacing the trial epicenter. Angular displacement δy northward.

The partial derivatives thus become

$$\frac{\partial \Delta}{\partial x} = - \sin \zeta_o \cdot 0.0090 \text{ deg/km} \tag{7}$$

$$\frac{\partial \zeta}{\partial x} = -(\cos \zeta_o / \sin \Delta) \cdot 0.0090 \text{ deg/km} \tag{8}$$

$$\frac{\partial \Delta}{\partial y} = - \cos \zeta_o \cdot 0.0090 \text{ deg/km} \tag{9}$$

$$\frac{\partial \zeta}{\partial y} = (\sin \zeta_o / \sin \Delta) \cdot 0.0090 \text{ deg/km}. \tag{10}$$

Equations (7) and (9) are the conventional relations for evaluating $\partial t/\partial x = (\partial t/\partial \Delta) \cdot (\partial \Delta/\partial x)$ and $\partial t/\partial y = (\partial t/\partial \Delta) \cdot (\partial \Delta/\partial y)$ in equation (1); equations (8) and (10) are the necessary relations for implementing the new feature of the location program, equation (2).

We see that the revised location algorithm treats back azimuth residuals in the same manner as it does arrival time residuals. This approach is consistent with the method which was already employed by the program LOC for handling measurements, made at arrays, of back azimuth and slowness (Julian, 1973). Comparing equations (1) and (2) shows that the epicentral coordinates x and y may be significantly better determined than the depth h and origin time T if the data base consists of only a few observations of arrival time and many observations of back azimuth. We note that this situation is unlikely to arise in practice and, even if it should happen, no conceptual difficulties are presented by it since the uncertainty in the determination of each of the four coordinates may be estimated by computing the standard errors and the four-dimensional error ellipsoid (Flinn, 1965, 1969; Evernden, 1969).

There is one additional feature of the location algorithm which ought to be mentioned since it subsequently will be shown to have an important effect on the results. The program LOC does not treat all data equally; instead, each arrival time residual is assigned a weight w_i so that some measurements can influence the calculations more strongly than can others. These weights are chosen to be the reciprocals of the standard deviations which are estimated *a priori* for each measurement, and hence they are different for, say, P_n arrival times than for those of teleseismic P. At the time of program execution one may invoke an option which causes the weights to be adjusted at the end of each iteration so that the influence of outlying data points is diminished.

It has been the authors' experience with the use of arrival time data, however, that invoking this program option is rather ineffective in decreasing the absolute error of location. The algorithm for the computation of new weights at the end of each iteration is thus in need of improvement, so the LOC program is customarily run with the *a priori* weights left unchanged. The revised LOC program handles the weighting of back azimuth data in the same manner as the previous version handled the weighting of arrival time data. Because the weights are the reciprocals of the standard deviations of the measurements, the residuals become dimensionless quantities, and arrival time data and back azimuth data can then be processed in the same fashion, the only difference being the size of the weights. We shall discuss this similarity further when we examine the computation of confidence intervals.

In analogy to the treatment of arrival time data, a certain standard deviation, which we take to be the same for every observation, is assigned *a priori* to each measurement of back azimuth, and the program may then be run with the weights fixed or adjustable at the end of each iteration. In light of our previous experience with arrival time data, we chose to run the program with the *a priori* weights left unchanged. Running the program with the weights fixed, however, results in a location which is too heavily influenced by outlying (and hence probably erroneous) values of the residuals. We have attempted to alleviate this situation by means of an *ad hoc* change in the program logic. Arrival time residuals, however large, are processed as before, but back azimuth residuals are assigned weight zero if they exceed some threshold, arbitrarily set to 15°. An exception is made in the case of residuals greater than 165°, since they are assumed to occur only if the surface-wave particle-motion, assumed retrograde, was actually prograde. In this case the back azimuth is corrected by 180° and the weight remains unchanged. This search for anomalous back azimuth residuals takes place independently during each iteration, so borderline values may alternately be included or be dropped as the trial epicenter changes from iteration to iteration. This procedure is far from optimum, and it would be preferable to replace it by one which assigns a weight which gradually decreases with increasing residuals. As we have said, however, unsatisfactory results were obtained with the pre-existing program option designed to implement such a weighting scheme, so for the purposes of this report we have chosen to use instead the *ad hoc* weighting scheme. The consequences of this choice will be discussed when the results of an evaluation of the algorithm are presented.

PROGRAM DOCUMENTATION CHANGES

Users of LOC should be aware of two changes which were made to the program code in order to implement the new algorithm described in the previous section of this report. The first change is that the program now ignores the variable IWAIT, a switch which indicates whether new weights should be assigned to all residuals at the end of each iteration. This change was made in order to implement the *ad hoc* weighting scheme described previously. The value which is input for IWAIT may thus be set to either zero or one without affecting the computations. The second change involves the input of the data measurements, which are now entered in the following format:

STA - station name

TABL - travel-time table; 'HEbb' for Herrin; 'JBbb' for J-B

PHAS - seismic phase; 'Pbbbbbbb', 'P.Nbbbbb', 'P.Gbbbbb', etc.

NNHR - arrival time hours

NMIN - arrival time minutes

ITN - 10.* arrival time seconds

SIGNL - 'IPbb' for impulsive P, 'EPbb' for emergent P

AZM - back azimuth in degrees

VEL - measurement of $dt/d\Delta$ made at an array. It appears that the LOC program ignores the value which is read in for this parameter.

AZCODE - a character string to identify back azimuth measurements made by means of Smart's (1978) three-component processor at a single station rather than by means of beam steering at an array. This character string is 'AZOK' for measurements which the analyst deems to be of good quality (i.e., for which the Smart processor yields a large F-statistic) or 'AZBD' for measurements in which the analyst has less confidence. The former measurements are assigned *a priori* standard deviations of seven degrees, in accordance with the average residuals measured in regions of "good" L_g propagation (Smart, personal communication); in analogy to 'IP' and 'EP', "bad" back azimuth measurements are assigned *a priori* standard deviations twice as large as are the "good" ones, i.e., $\sigma_{\text{bad azimuth}} = 14^\circ$.

The format for the inputs of these ten variables is:

(A8, A4, A8, 2X, 2I3, I4, 4X, A4, F5.1, F5.2, 1X, A4).

Only the first seven variables are needed for arrival time measurements, and only the first, eighth, and tenth are needed for back azimuth measurements made with the Smart processor.

TEST OF THE ALGORITHM

In order to test the new algorithm, we have run twenty-four trials of the program using various combinations of arrival time and back azimuth measurements. These twenty-four trials are listed in Table I, and they will subsequently be described individually. The tests were performed using data recorded from the nuclear explosions SALMON (22 Oct. 1964, near Hattiesburg, Miss., 5.3 kt) and GNOME (10 Dec. 1961, near Carlsbad, N. Mex., 3.1 kt). Back azimuth measurements for these two events have previously been used for event location by Smart (1978). Table II presents the back azimuths and P-wave arrival times measured for SALMON. For this event we have used four stations which were not included in Smart's data set, and we have supplemented the back azimuth measurements with arrival time measurements at all stations except one. SALMON was of course recorded at many more stations than the twelve which are considered in this report; it was at these twelve, however, that three-component short-period digital records of L_g were available from which the back azimuth could be determined using Smart's (1978) surface-wave processor. In order to simplify the interpretation of the results, all ambiguities of 180° introduced by incorrect assumptions of prograde/retrograde motion were resolved *a priori* when the data base was created. For the case at hand, this *a priori* correction was unnecessary, since the program automatically adjusts *a posteriori* the back azimuths by 180° if it is necessary to do so in order to achieve consistency with the trial epicenter, and since an accurate trial epicenter (in fact the true value) was known in this case. Although the *a priori* adjustments are unnecessary for this particular data set, we shall discuss later in this report circumstances under which these adjustments might be needed because the *a posteriori* adjustments could fail. The data base for GNOME (Table III) consisted of the seven back azimuth measurements used by Smart (1978), supplemented by corresponding arrival time measurements. The location of the stations and the directions of the measured back azimuths are shown in Figure 2a for SALMON and Figure 2b for GNOME. Note that the arrows which represent the measured back azimuth in these figures should not be expected to intersect at the events' epicenter even if the measurements were perfect, since they are drawn as straight lines rather than as great circle arcs.

We shall now analyze on a case-by-case basis the results of the twenty-four tests listed in Table I as they were applied to the data base for SALMON. The results of these tests are listed in Table IV and are shown in Figure 3a-s. The first of the twenty-four tests used all the measurements of arrival time and back

TABLE I

Description of the Tests of the Location Algorithm

Trial #	Arrival Times	Back Azimuths	Comments
1	Yes	Yes	all stations; depth free
2	Yes	Yes	all stations; depth restrained
3	Yes	No	all stations; depth free
4	Yes	No	all stations; depth restrained
5	No	Yes	all stations
6	one station	all stations	
7	Yes	No	3 "distant" stations
8	three stations	all stations	3 "distant" stations
9	Yes	No	3 "near" stations
10	three stations	all stations	3 "near" stations
11	Yes	Yes	3 "distant" stations
12	Yes	Yes	3 "near" stations
13	Yes	Yes	stations in one sector
14	Yes	No	stations in one sector
15	No	Yes	stations in one sector
16	Yes	Yes	2 "distant" stations
17	Yes	Yes	2 "near" stations
18	No	Yes	3 "distant" stations
19	No	Yes	3 "near" stations
20	Yes	Yes	Smart's network
21	Yes	No	Smart's network
22	No	Yes	Smart's network
23	Yes	Yes	all stations; poor initial location
24	No	Yes	all stations; poor initial location

TABLE II

P-Wave Arrival Time and Back Azimuth Measurements for SALMON

Station	Distance (degrees)	Forward Azimuth	Back Azimuth	Measured Arrival Time	Measured Back Azimuth
EU-AL[1]	2.178	41.0	221.9	16:00:36.5	204.2
JE-LA[1]	2.185	287.7	106.4	00:37.4	100.8
BL-WV	9.514	43.5	228.2	02:18.5	235.7
VO-IO	11.249	350.2	168.7	02:40.6	182.1
BR-PA	12.364	41.9	228.2	02:54.0	208.2
DH-NY	16.143	42.6	231.4	03:45.3	225.0
GP-MN	16.772	350.8	168.3	03:51.4	167.5
LS-NH	19.078	41.8	232.7	—	222.4
RK-ON	19.914	352.3	169.6	04:30.5	174.5
KN-UT[1]	20.115	293.2	100.1	04:37.3	91.3
HN-ME	22.460	42.0	235.7	05:00.0	243.3
EK-NV[1]	22.777	297.9	102.7	05:05.4	95.0

[1] not included in the data set of Smart (1978)

TABLE III

P-Wave Arrival Time and Back Azimuth Measurements for GNOME

Station	Distance (degrees)	Forward Azimuth	Back Azimuth	Measured Arrival Time	Measured Back Azimuth
AM-OK	5.698	70.1	253.6	19:01:25.8	240.5
SJ-TX	6.685	132.5	315.2	01:39.7	328.3
MP-AR	9.257	72.5	258.4	02:12.9	248.5
WR-AR	11.136	66.7	253.8	02:37.6	257.0
CV-TN	14.110	71.1	260.4	03:17.5	263.7
BL-WV	19.262	67.1	260.1	04:25.4	264.2
DH-NY	25.019	58.3	256.0	05:24.8	257.6

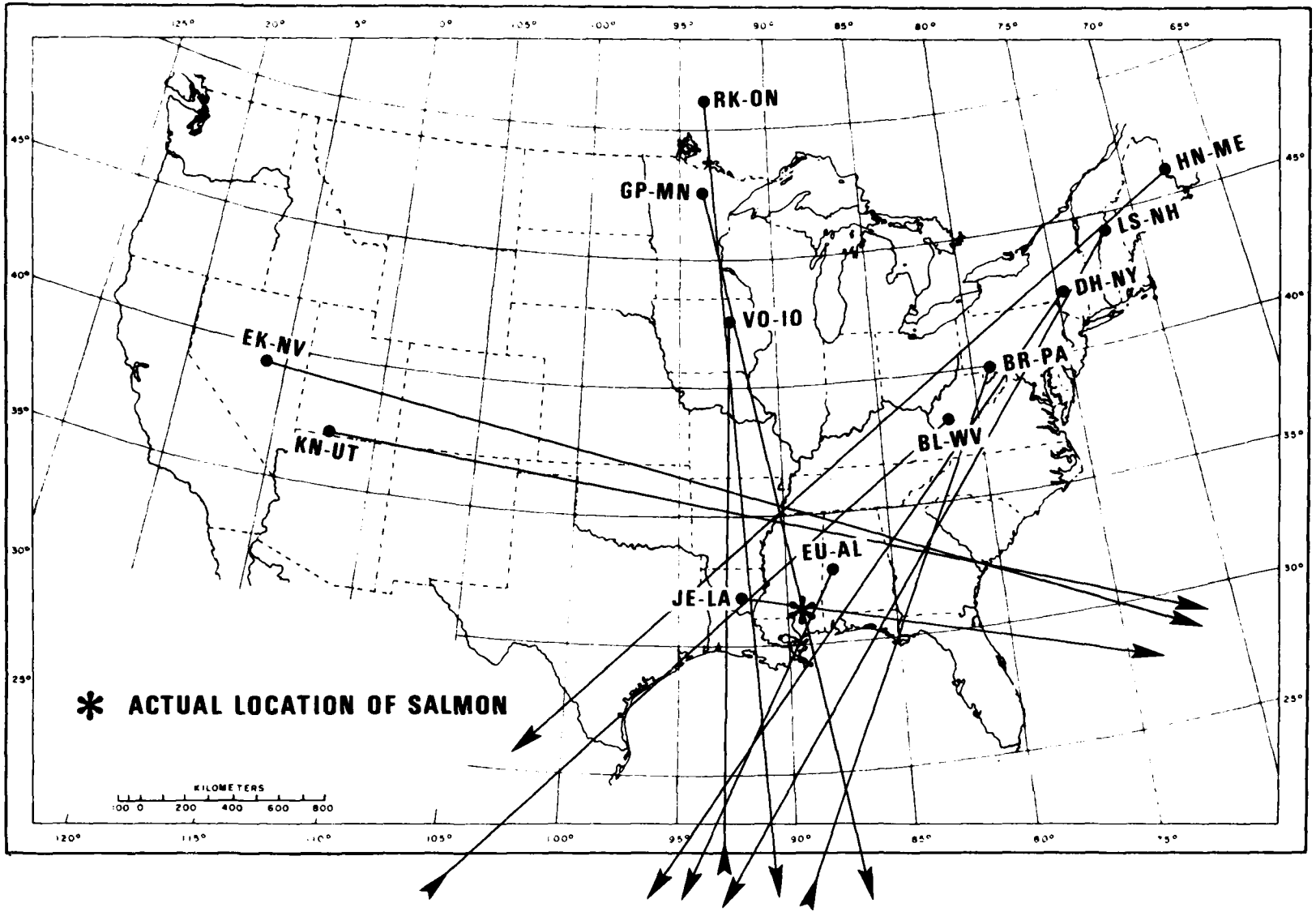


Figure 2a Back azimuths calculated using Smart's three-component processor on L_g. Reversed arrowhead indicates azimuth has been adjusted by 180°. SALMON (after Smart, 1980).

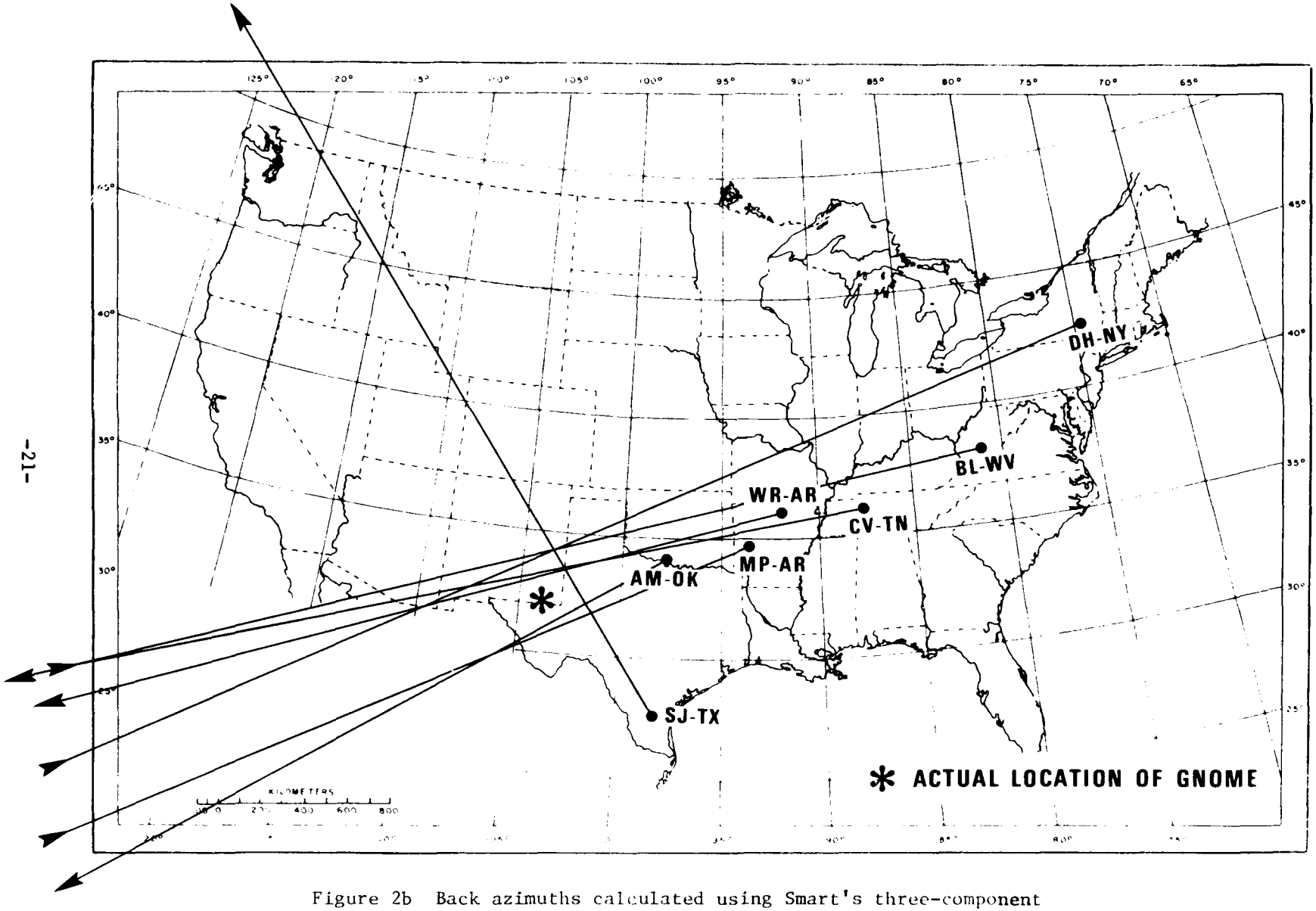


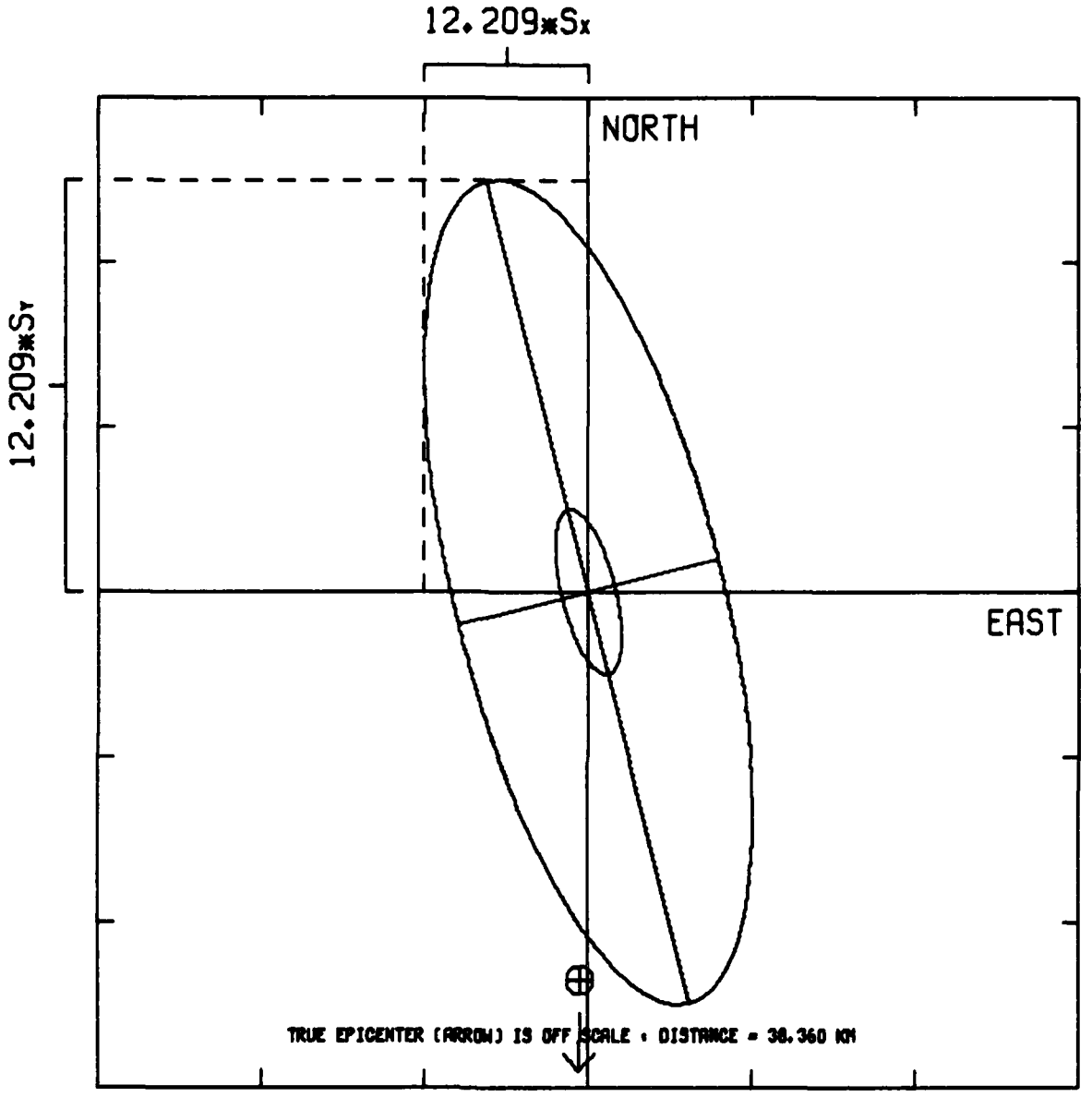
Figure 2b Back azimuths calculated using Smart's three-component processor on L_g. Reversed arrowhead indicates azimuth has been adjusted by 180°. GNOME (after Smart, 1980).

TABLE IV

Results of the Tests of the Location Algorithm Applied to SALMON

True epicenter and origin time: 31.142°N, 89.570°W, 16:00:00

Trial #	Lat (°N)	Long (°W)	Origin Time	Standard Errors (km) North	East	Absolute Error (km)	No. Azimuths Deleted	Reversed
1	no convergence							
2	31.488	89.562	16:00:02.7	2.34	0.94	38.36	2	
3	no convergence							
4	31.489	89.562	16:00:02.7	2.35	0.94	38.47		
5	31.130	88.807	16:00:00.0	34.23	28.08	72.76		
6	31.130	88.807	15:59:56.5	34.22	28.08	72.76		
7	31.504	89.511	16:00:03.5	645.14	122.04	40.52		
8	30.905	89.379	15:59:59.7	27.43	5.34	31.98		
9	no convergence							
10	no convergence							
11	32.950	89.839	16:00:12.4	104.60	20.63	202.04		
12	no convergence							
13	31.947	89.669	16:00:06.0	12.27	4.27	89.73	2	
14	31.951	89.670	16:00:06.0	12.26	4.27	90.18		
15	43.093	76.688	16:00:00.0	155.27	111.30	1748.01	3	2
16	34.531	90.190	16:00:21.7	222.54	42.42	954.61		
17	30.859	89.429	15:59:57.1	28.24	8.33	34.13		
18	34.481	90.290	16:00:00.0	171.15	345.62	376.31		
19	31.252	88.737	16:00:00.0	36.07	28.73	80.31		
20	31.451	89.589	16:00:02.1	8.10	2.84	34.38	1	
21	31.456	89.587	16:00:02.2	8.12	2.84	34.93		
22	27.671	91.242	16:00:00.0	234.83	130.68	418.26		
23	31.488	89.562	16:00:02.7	2.34	0.94	38.36	2	
24	31.527	90.467	16:00:00.0	23.76	81.12	95.42	2	

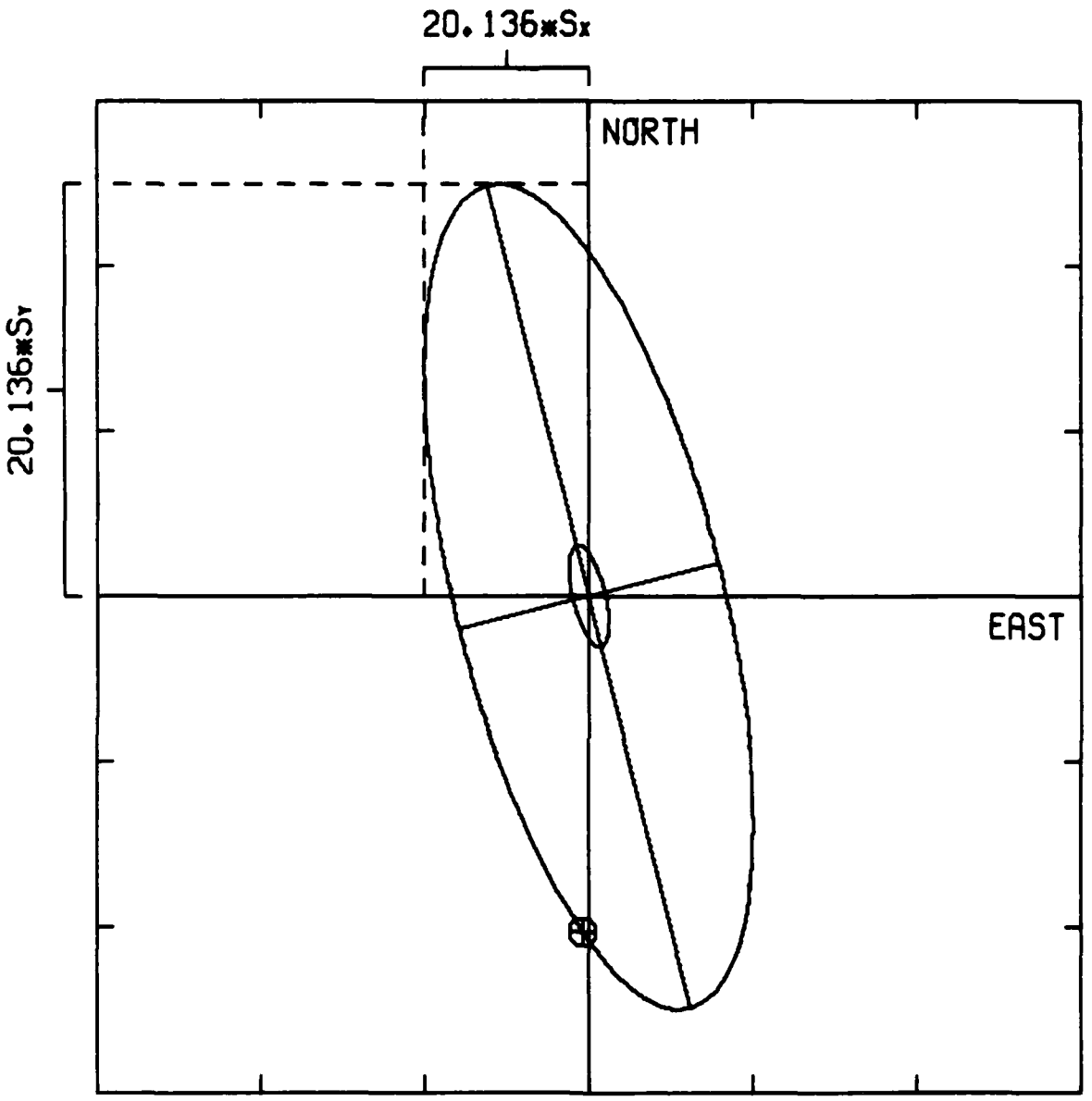


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 11.430 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3a Error Ellipses for SALMON, Trial 2

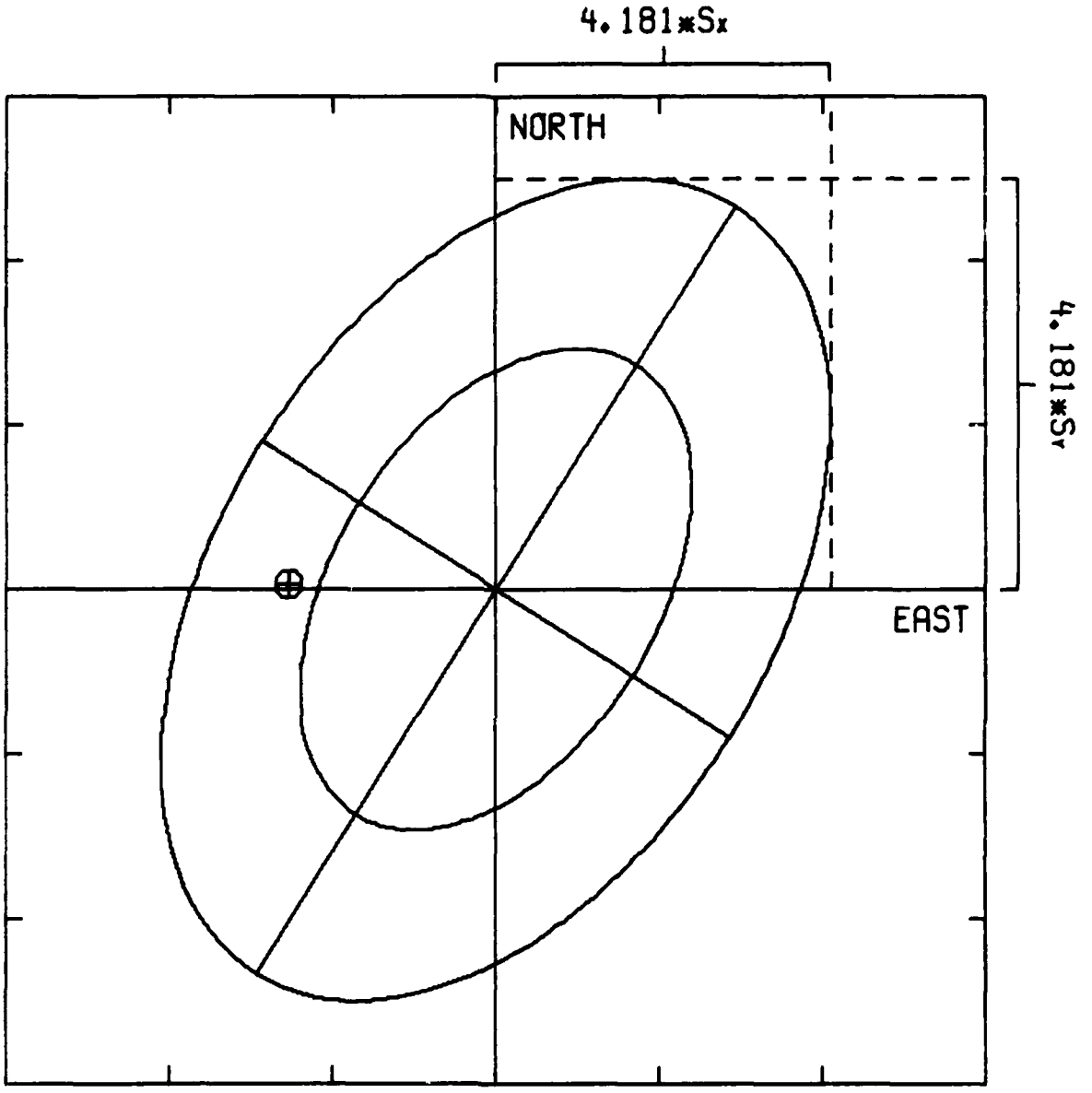


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
18.912 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3b Error Ellipses for SALMON, Trial 4

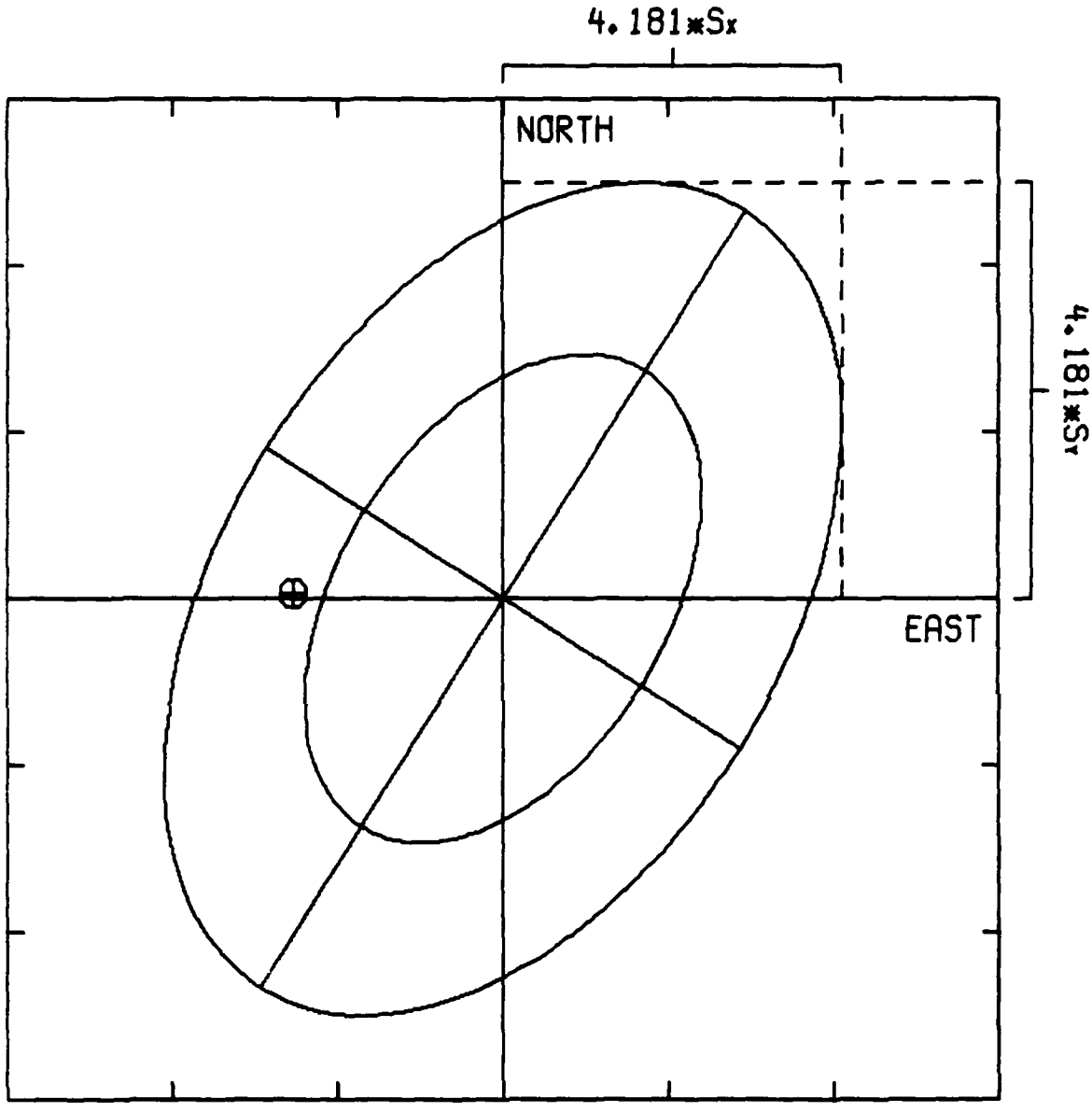


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
57.237 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3c Error Ellipses for SALMON, Trial 5



Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
57.235 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3d Error Ellipses for SALMON, Trial 6

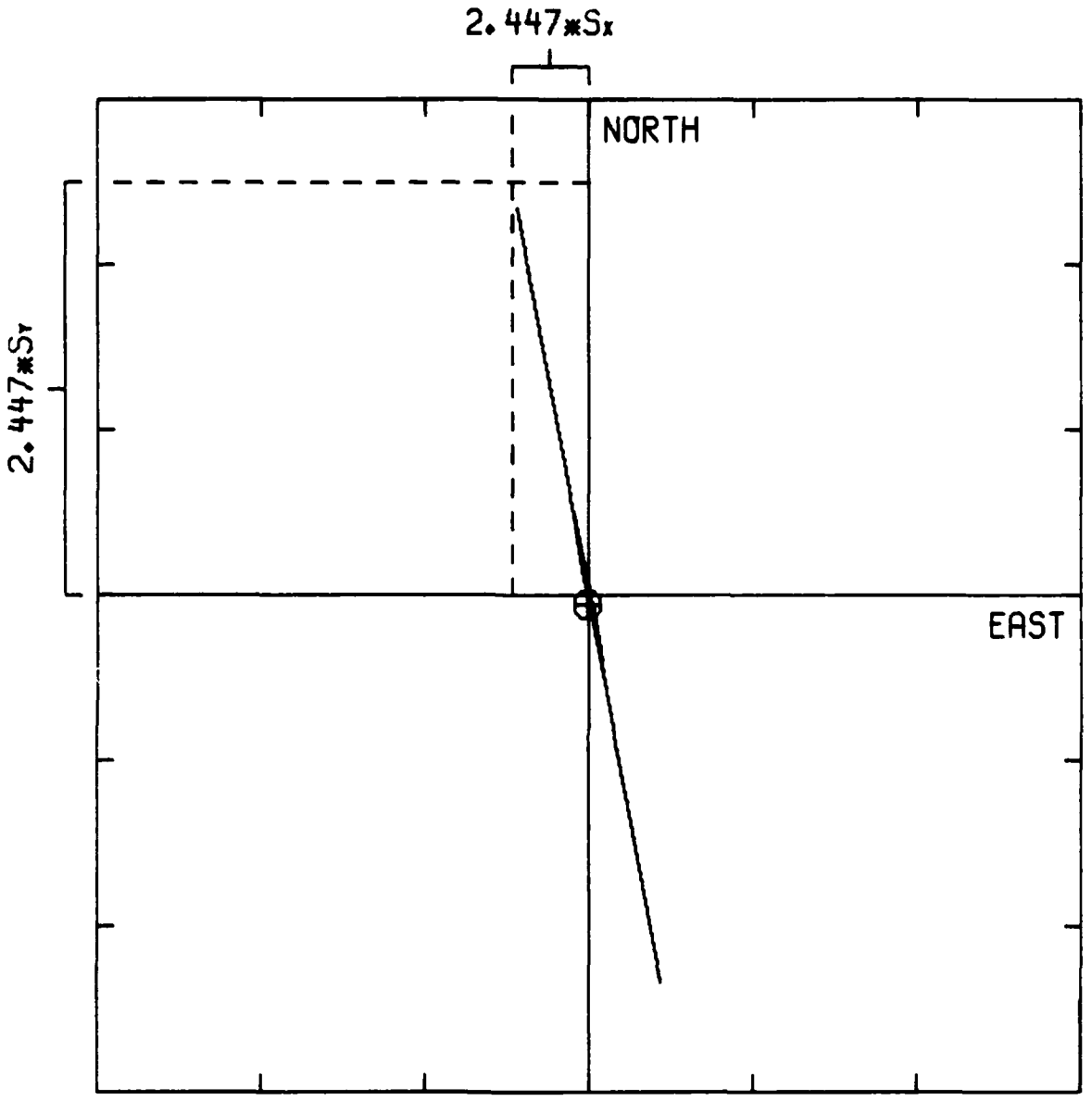
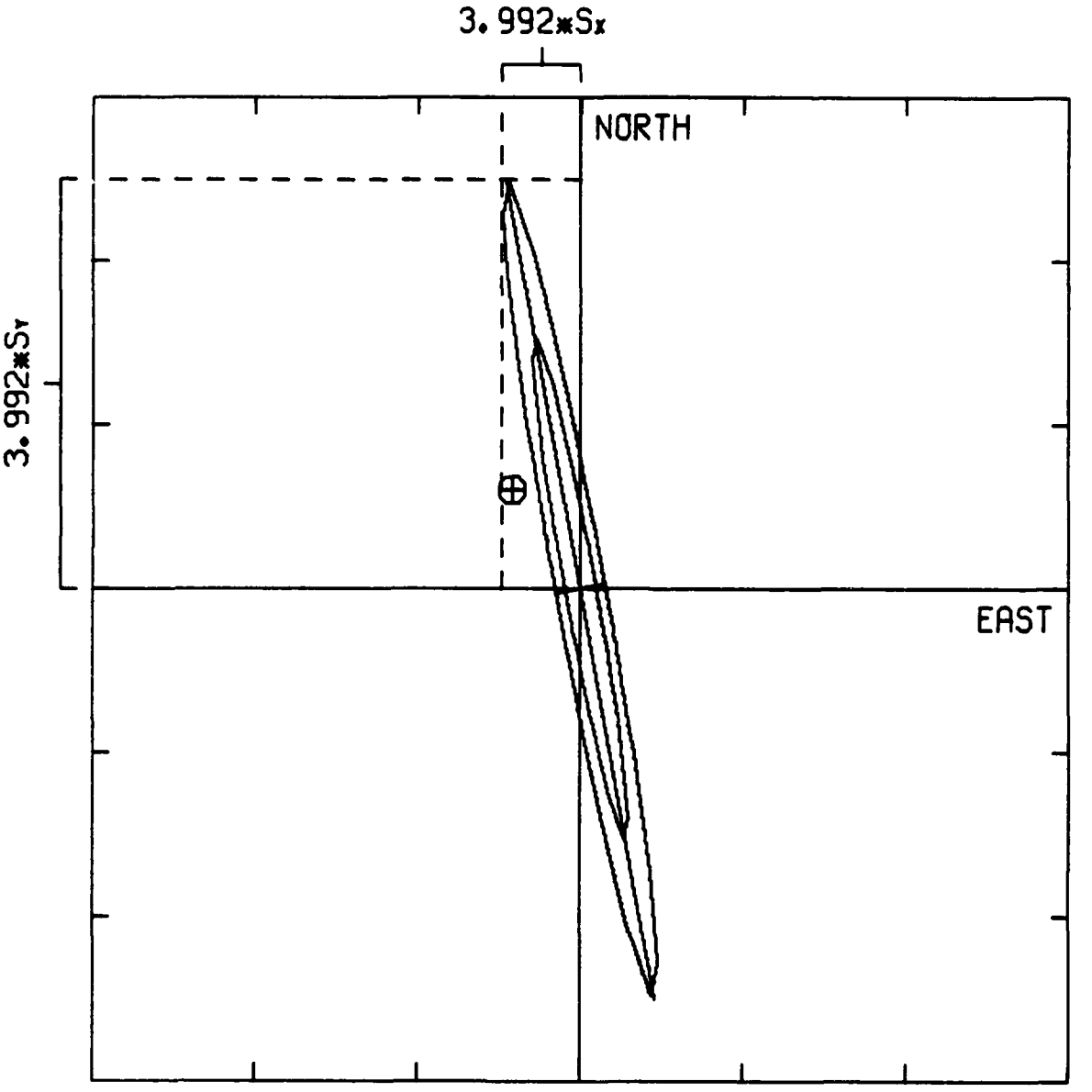


Figure 3e Error Ellipses for SALMON, Trial 7

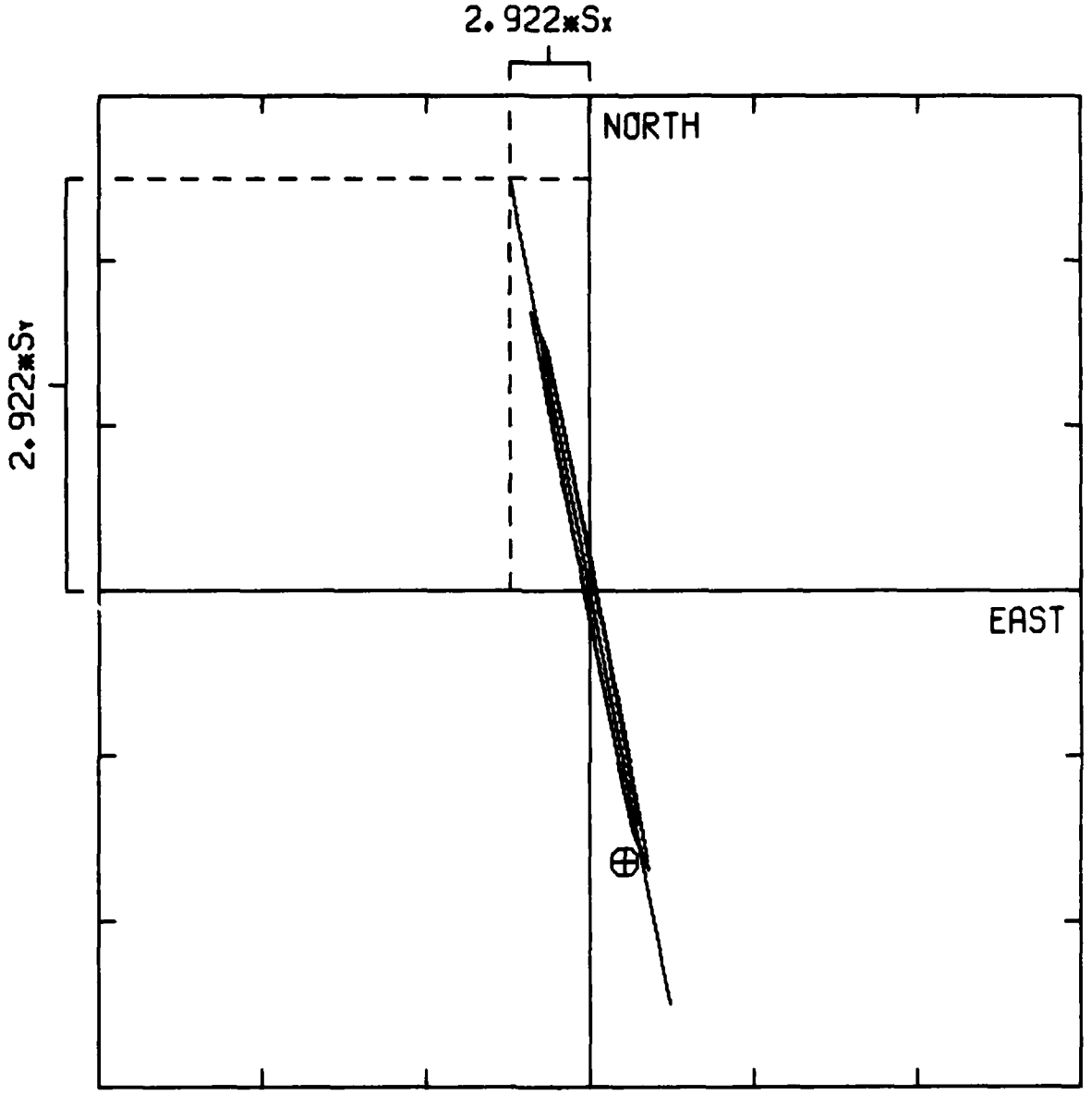


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
43.804 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3f Error Ellipses for SALMON, Trial 8

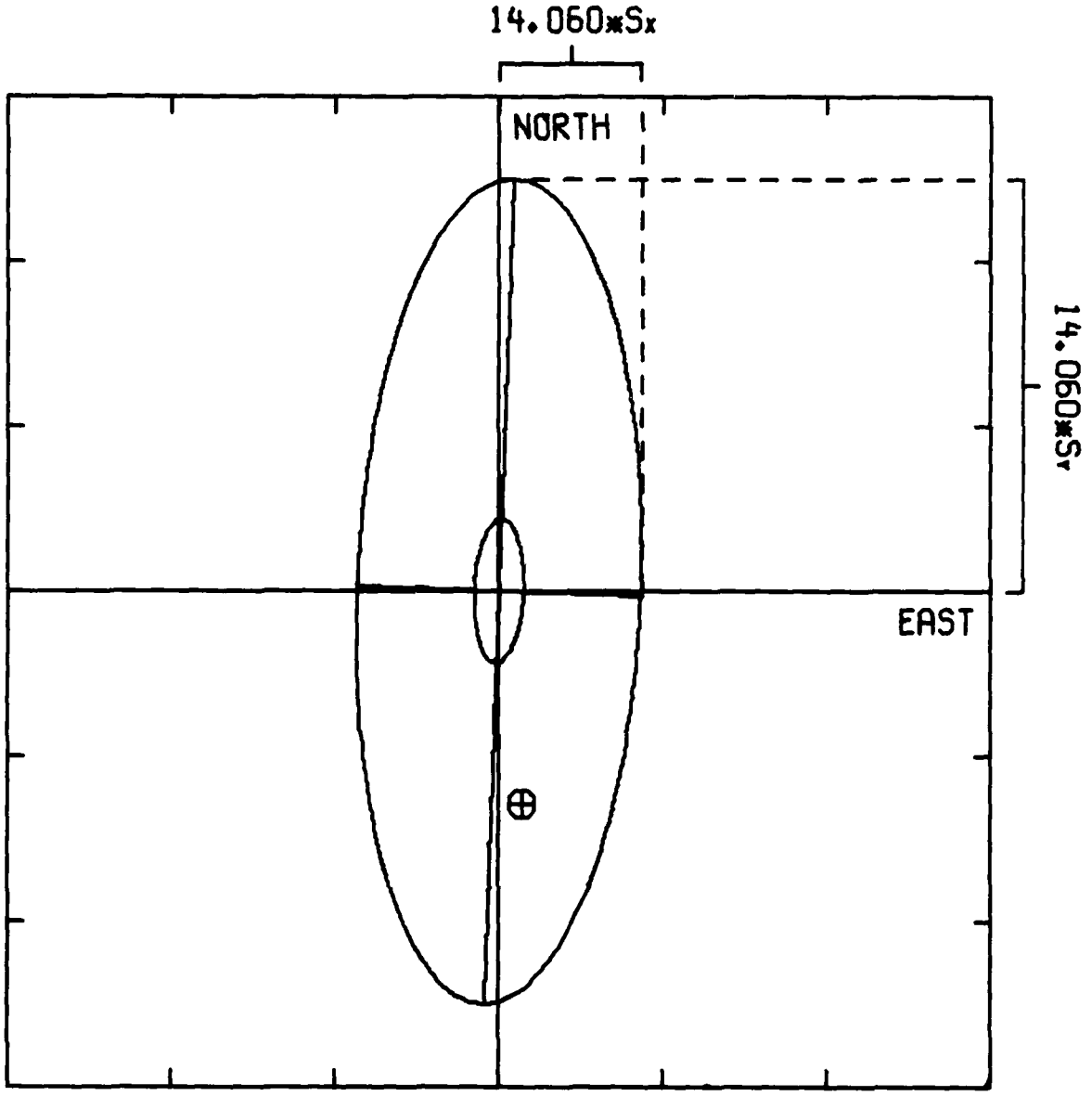


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
122.270 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3g Error Ellipses for SALMON, Trial 11

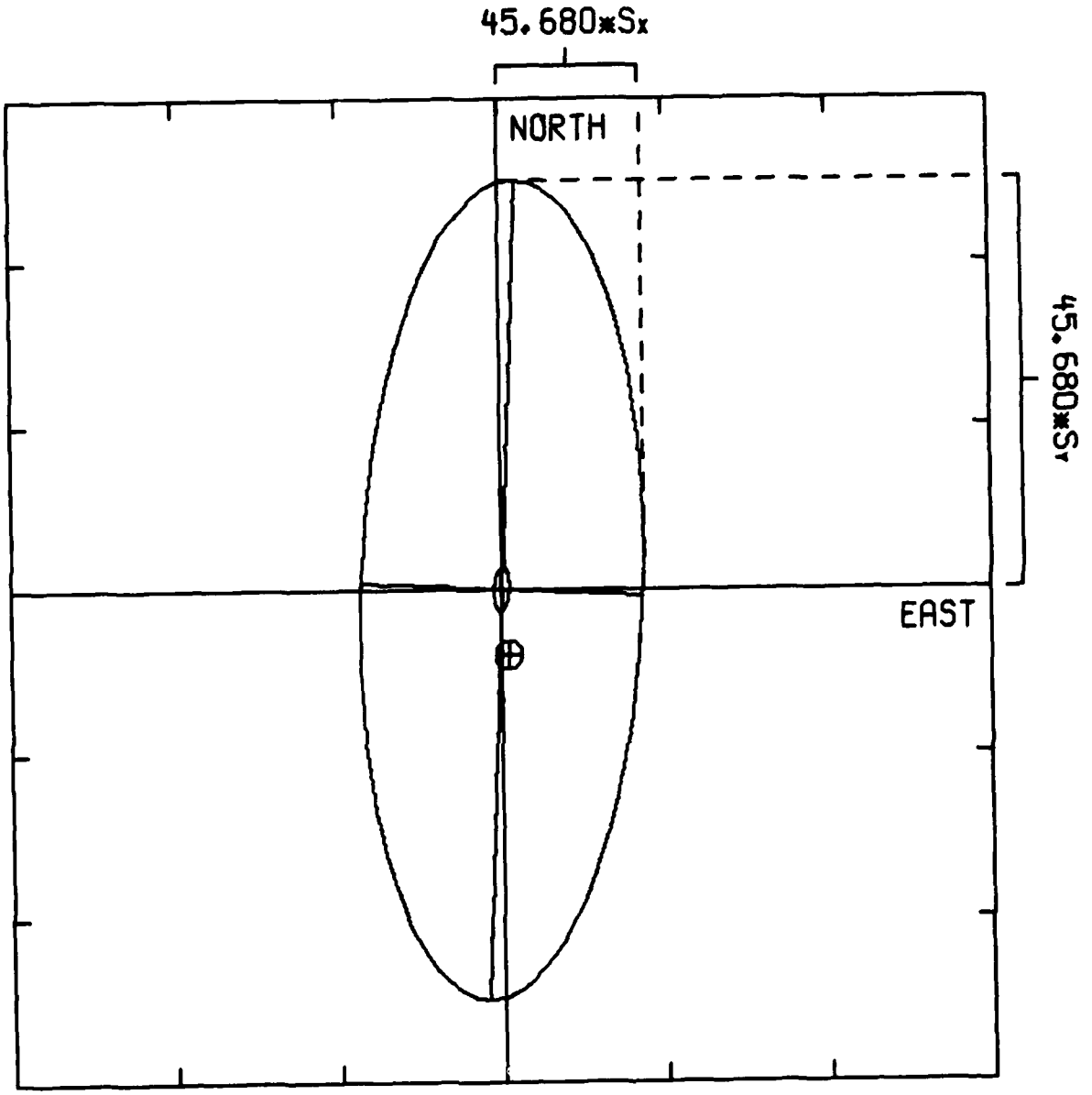


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
69.011 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3h Error Ellipses for SALMON, Trial 13

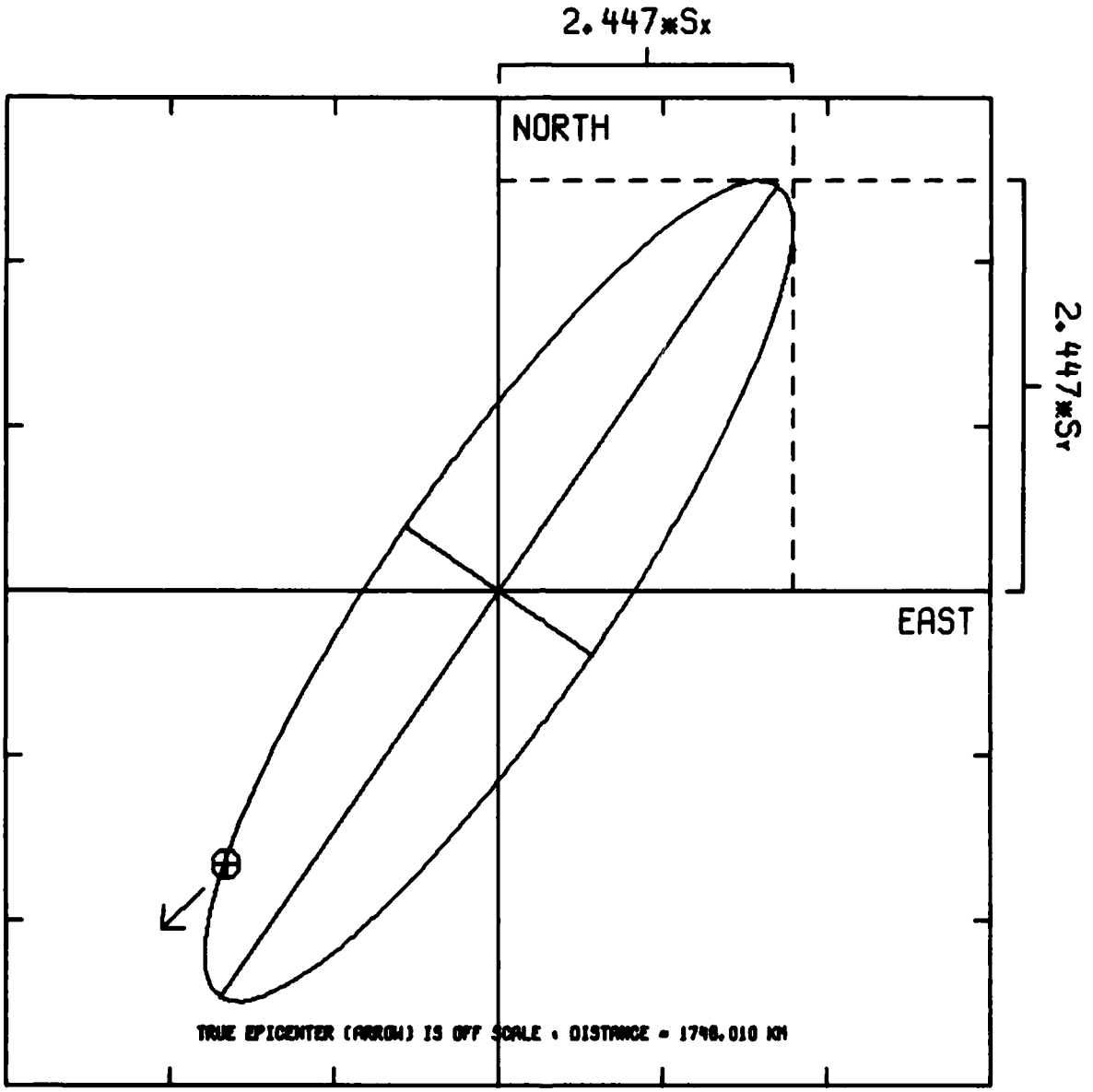


S_x : STANDARD ERROR (EAST)
S_y : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
224.053 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3i Error Ellipses for SALMON, Trial 14

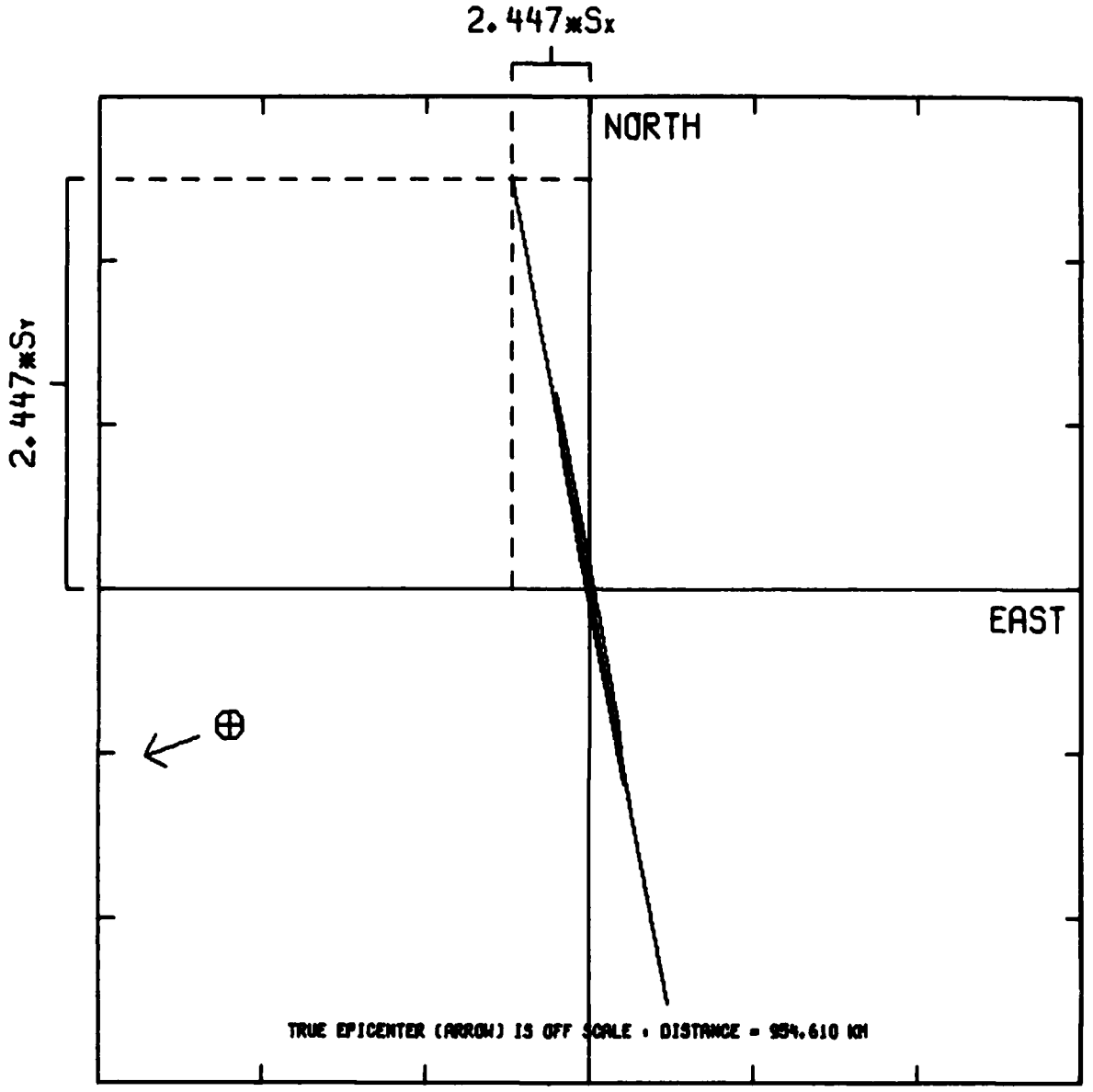


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 152.006 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 3j Error Ellipses for SALMON, Trial 15



Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
217.862 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 3k Error Ellipses for SALMON, Trial 16

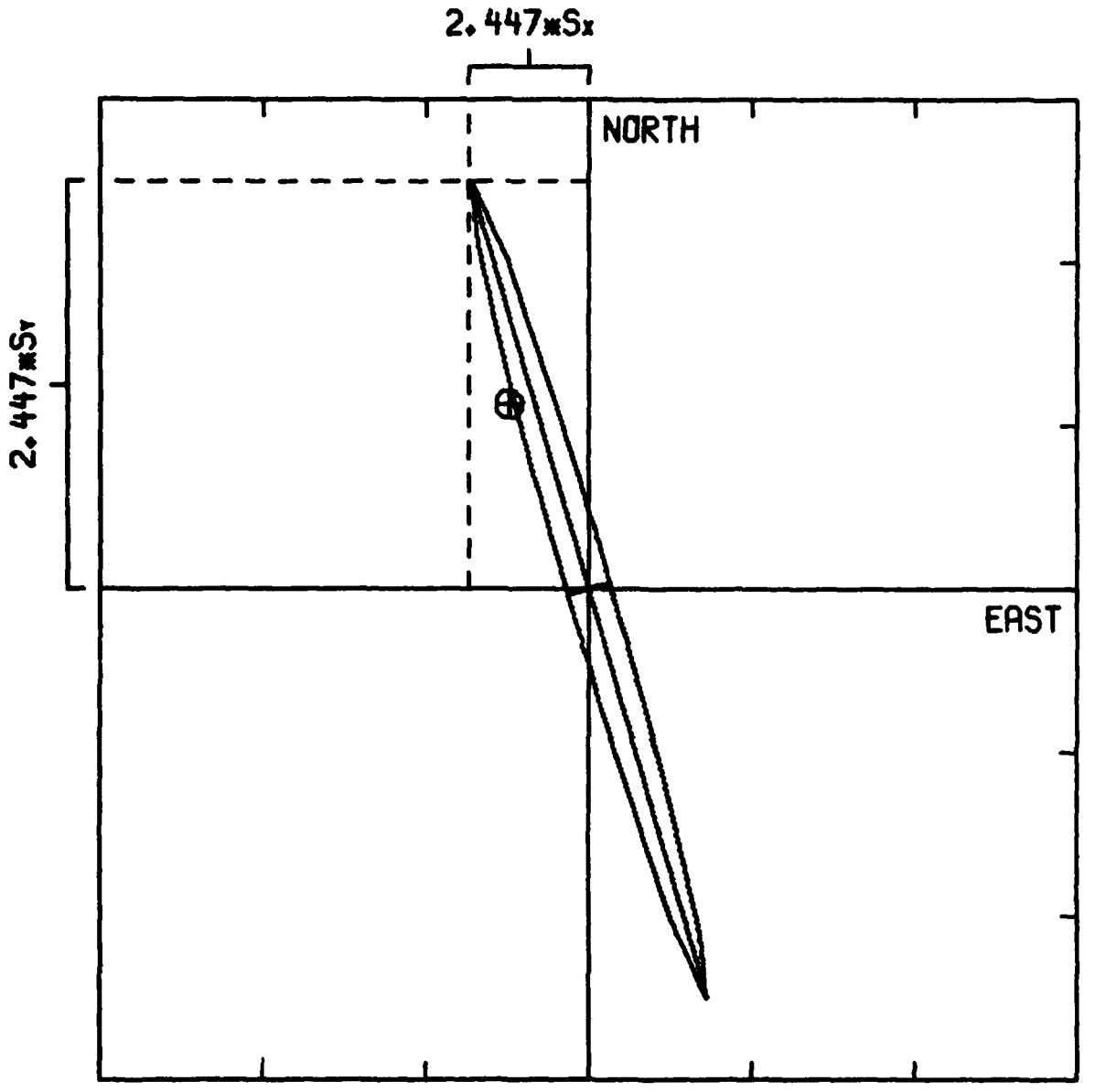


Figure 31 Error Ellipses for SALMON, Trial 17

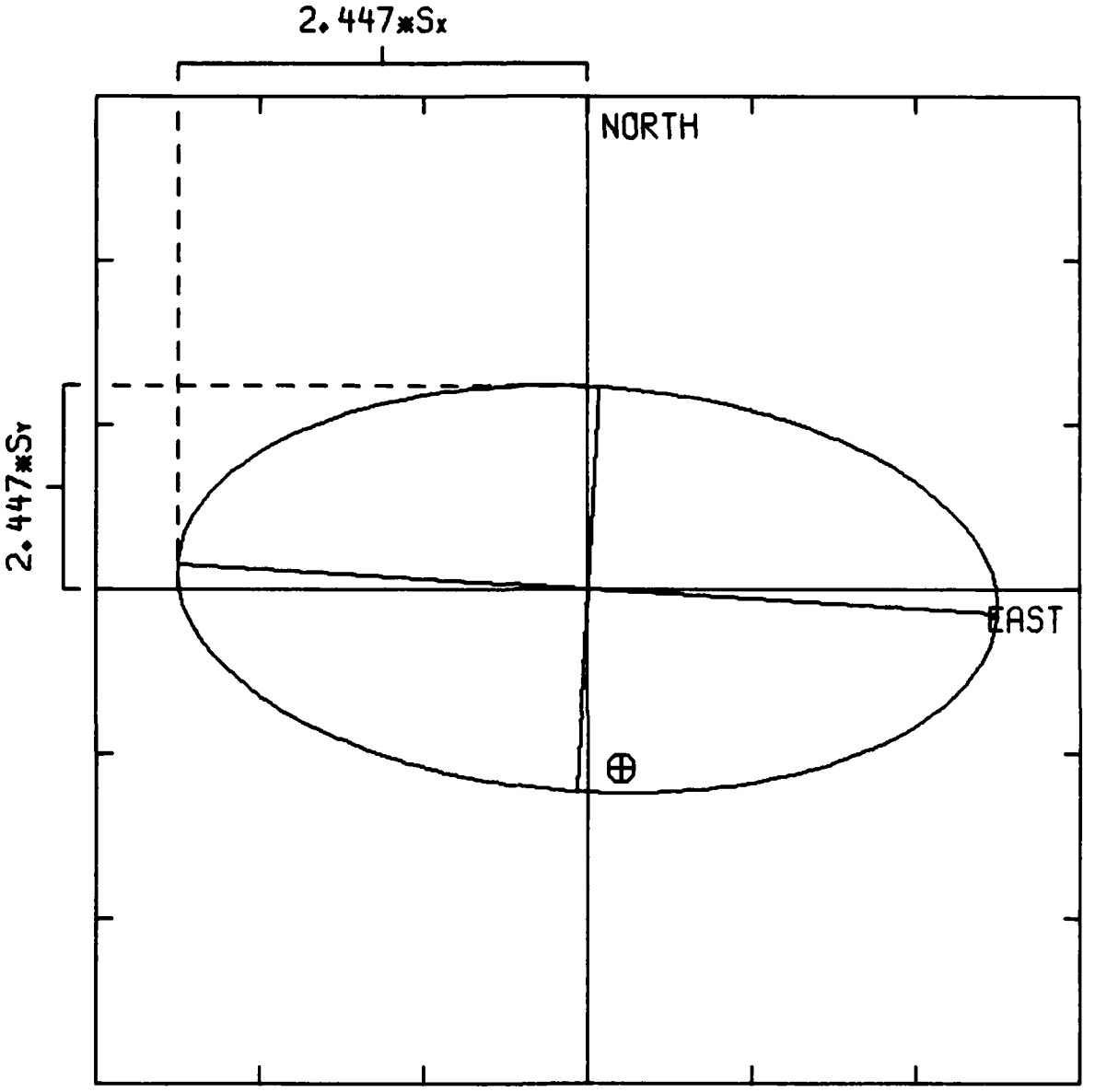
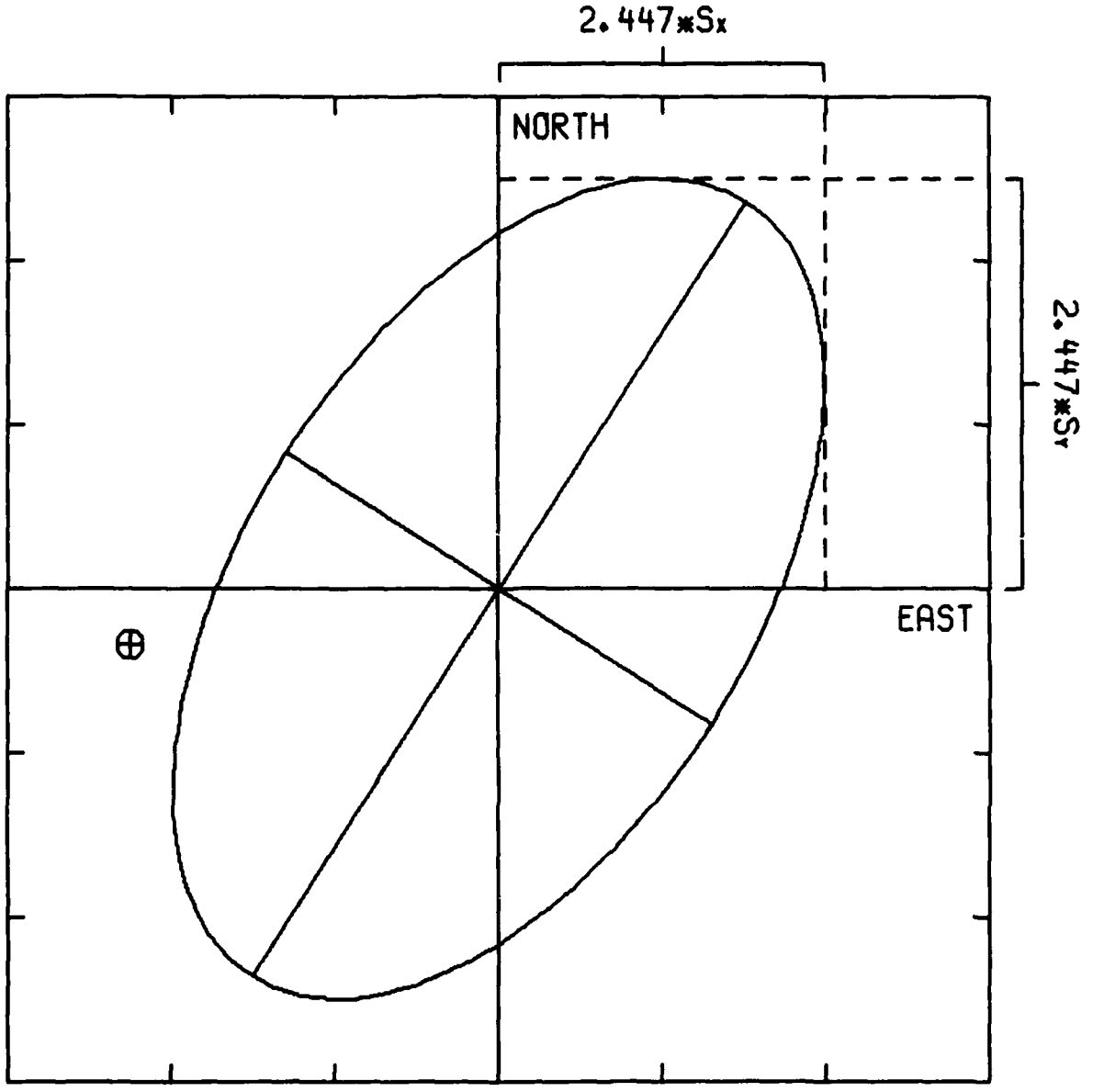


Figure 3m Error Ellipses for SALMON, Trial 18



Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
35.314 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 3n Error Ellipses for SALMON, Trial 19

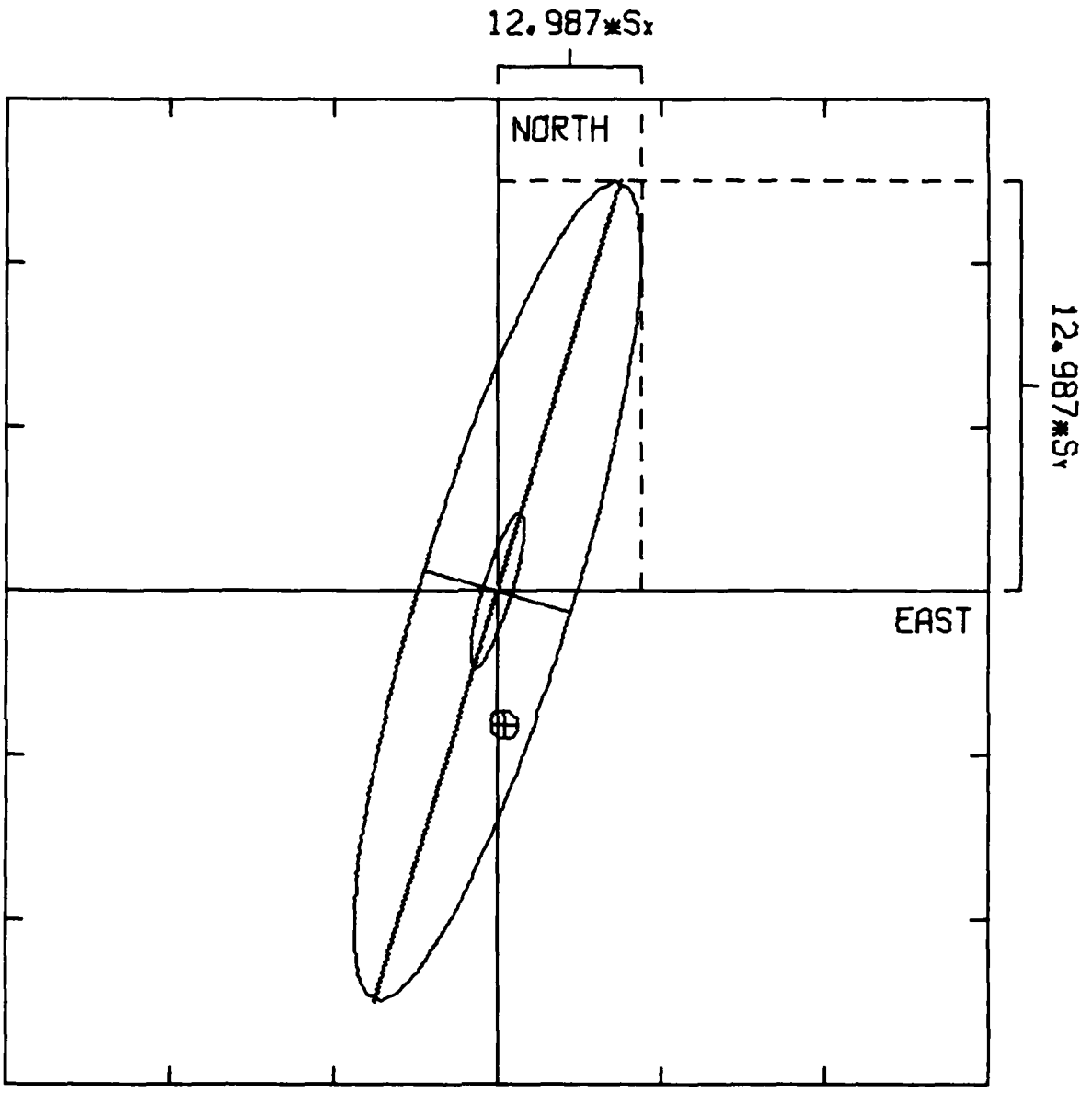
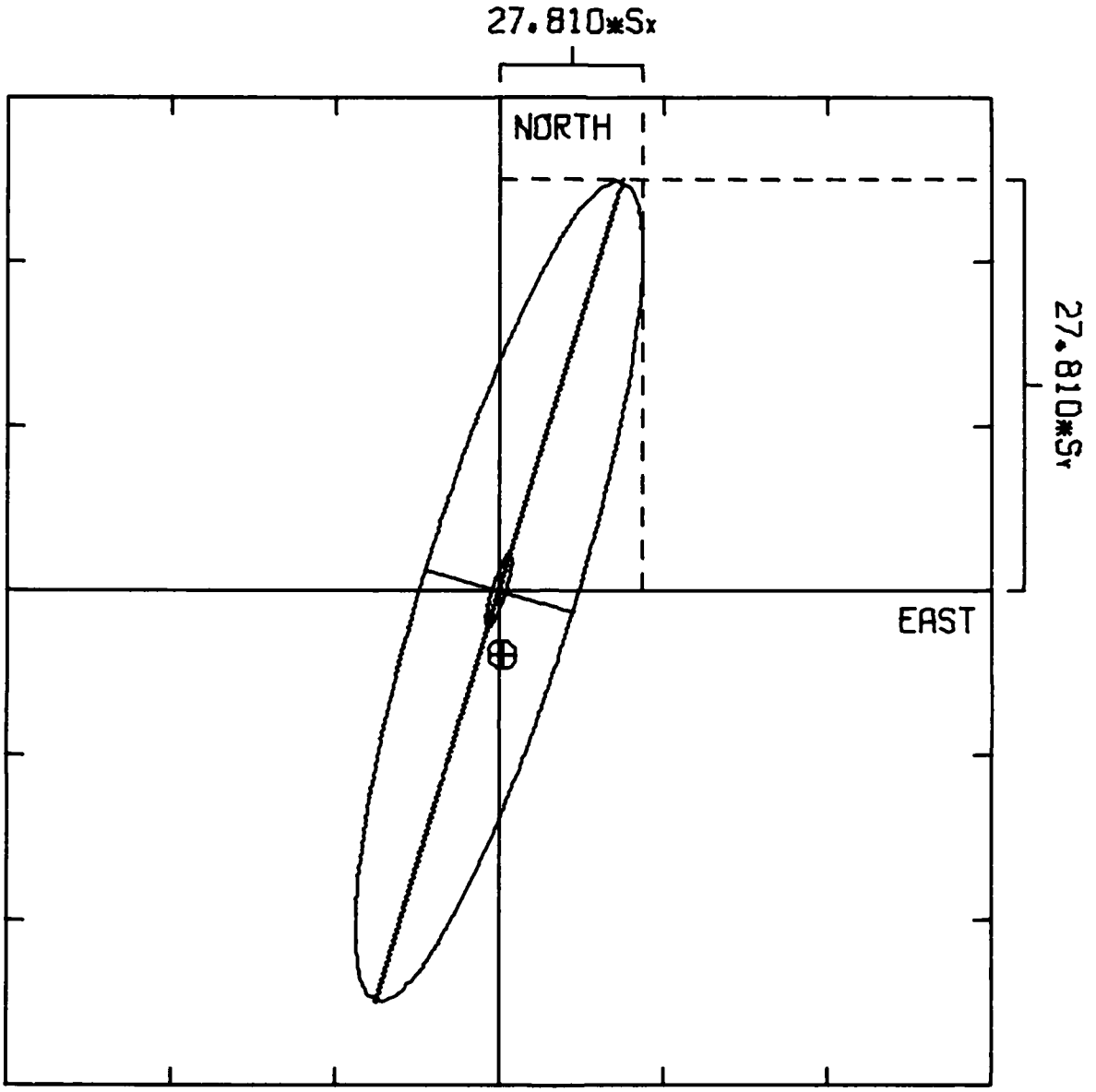


Figure 3o Error Ellipses for SALMON, Trial 20

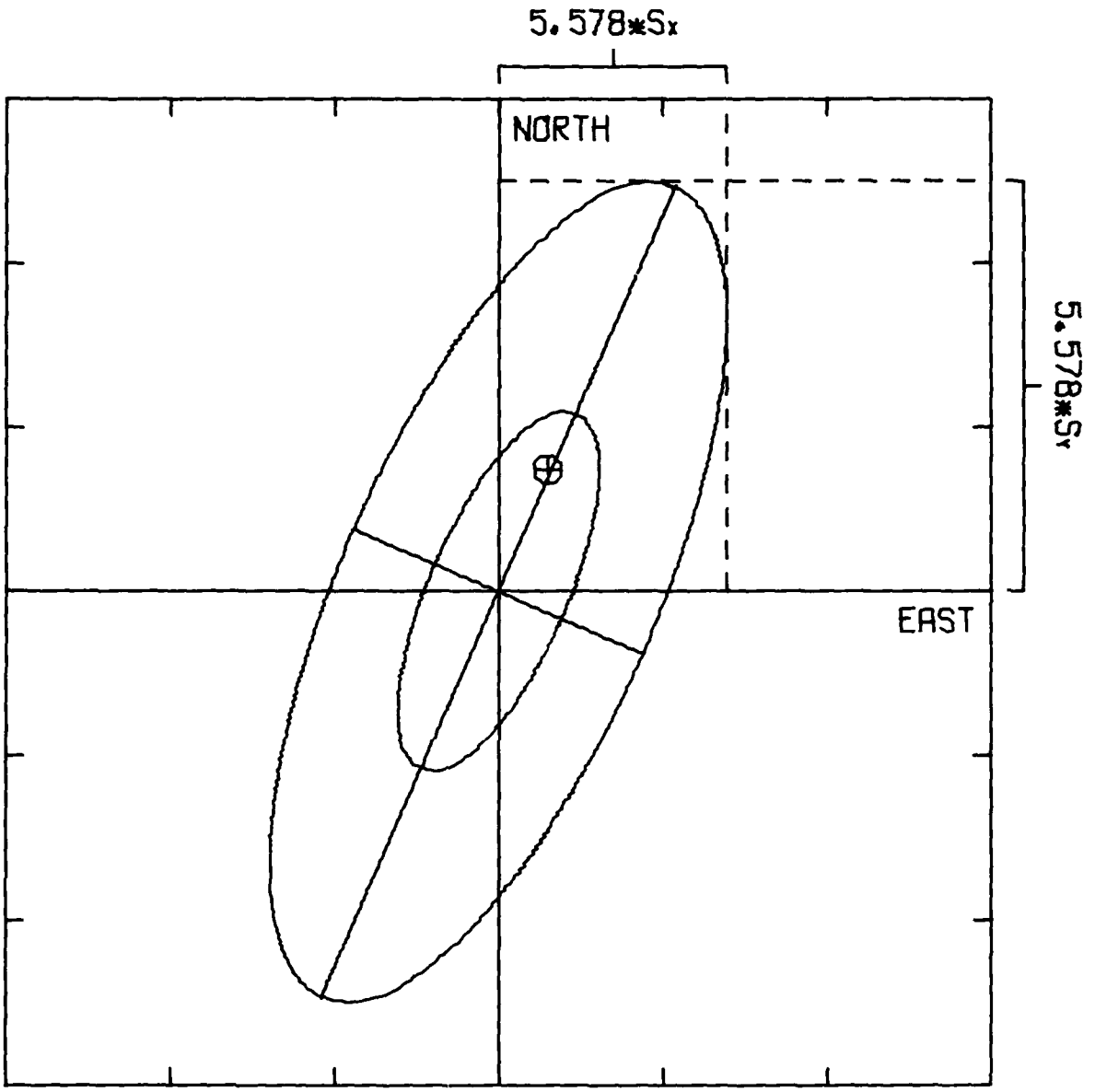


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 90.323 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3p Error Ellipses for SALMON, Trial 21

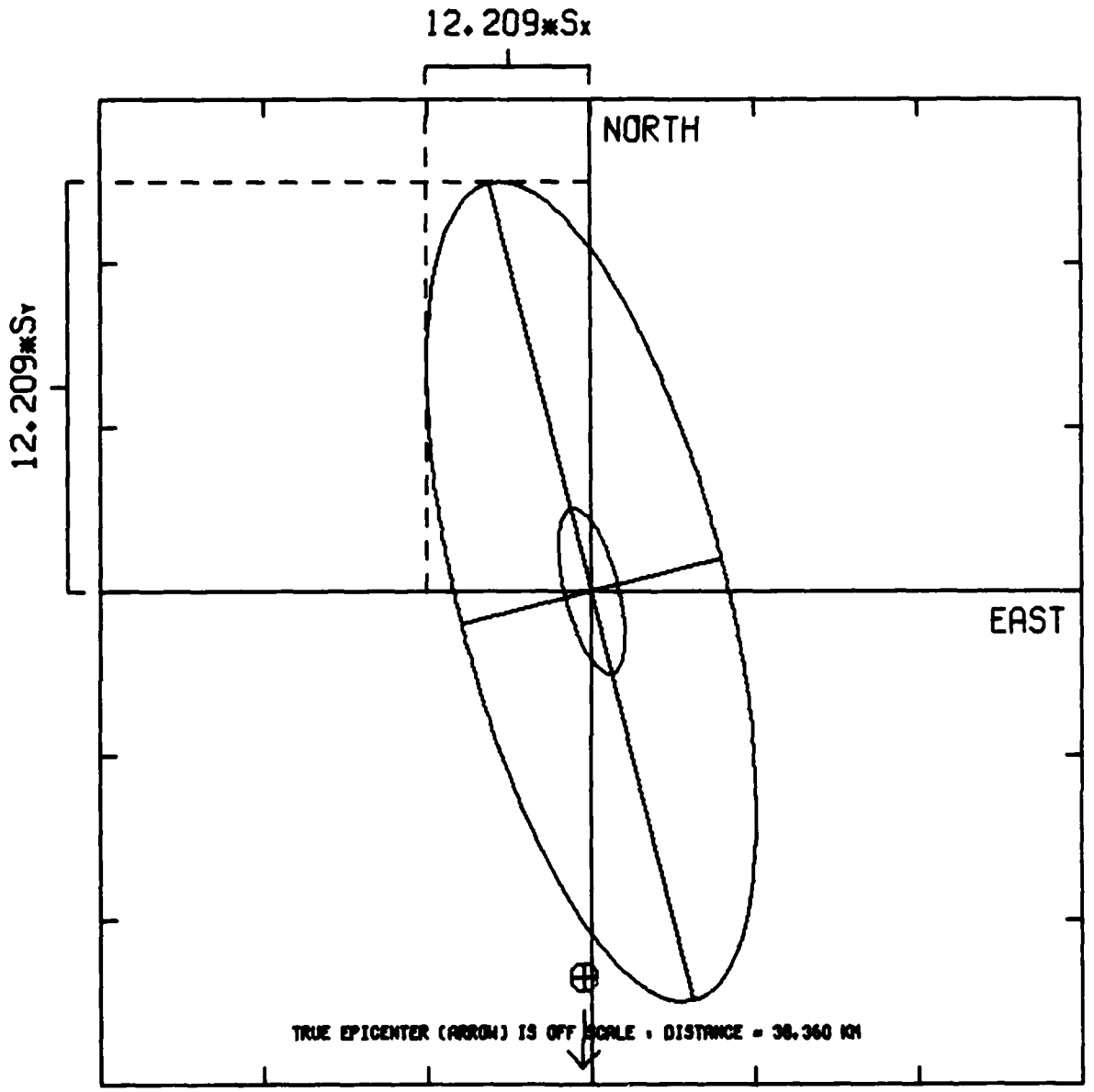


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
523.960 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3q Error Ellipses for SALMON, Trial 22

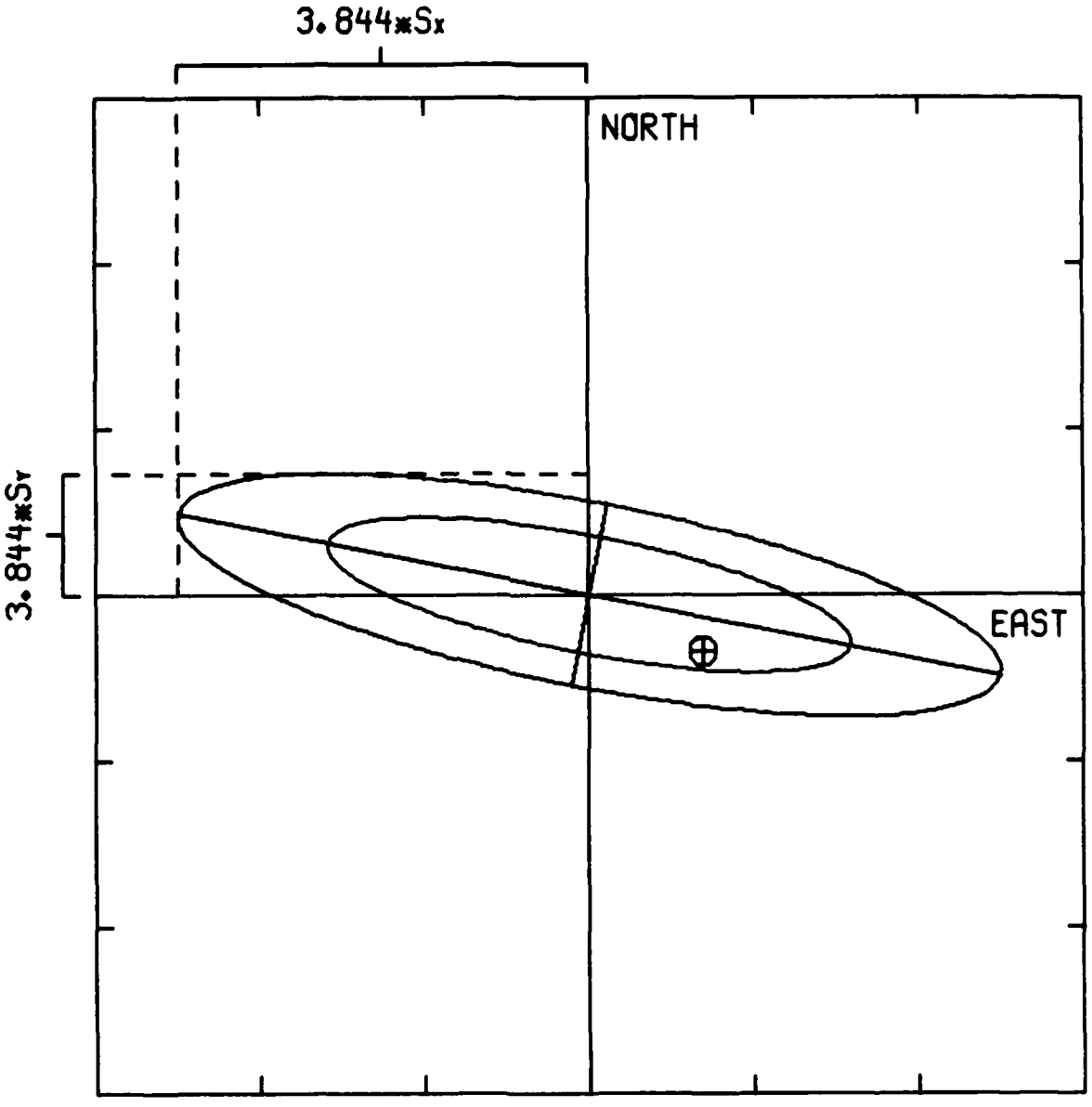


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 11.429 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3r Error Ellipses for SALMON, Trial 23



Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
124.744 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 3s Error Ellipses for SALMON, Trial 24

azimuths in the data base and made no restraint on the depth of the calculated hypocenter. For this case the iterative scheme failed to converge to a single location instead, the sum of the residuals alternately increased and decreased as the location found at the end of each iteration moved about, always at negative depth. When the depth was restrained to the surface, in the second of the twenty-four trials, the location rapidly converged to an epicenter 38.4 km distant from the true explosion site (cf. Table IV). Because we wished to see whether it was the algorithm using the back azimuth information which was causing the failure of the iterative scheme to converge in the depth-free case, we repeated these first two trials using only the arrival time measurement. These conventional runs of the LOC program are designated as Trials 3 and 4 in Table I. As is shown in Table IV, the iterative scheme in these trials showed the same behavior as in the first two trials which included back azimuth measurements, since the depth-restrained case converged and the depth-free case failed to do so. The inclusion of back azimuth information thus did not have a deleterious effect upon the location process; in fact, it had almost no effect at all, since the error in location using arrival-time measurements only (cf. Table IV) was 38.5 km, only 0.1 km greater than when the back-azimuth measurements were included. Figures 3a and 3b show that the calculated epicenters are scarcely different in the two cases (note from the legends that the scales on the two figures are different), but the confidence regions are smaller when the back-azimuth measurements are included in the data set. The confidence regions are in fact too small in this case; Figure 3a shows that they fail to include the true epicenter.

It is not surprising that such a small improvement in the location results from the addition of back-azimuth measurements to a data set consisting of several measurements of arrival time. An arrival time measurement which is accurate to within ± 1 sec defines the epicenters as lying on a circle which is centered at the station and has a radius of $\Delta \pm \delta$, where $\delta \approx 10$ km at $\Delta = 20°$. If several stations report arrival times, then, the location is given by the intersection of several annuli of thickness ~ 20 km, and the epicenter is thus determined to lie within a small region. A measurement of back azimuth which is accurate to within ± 7°, on the other hand, defines the epicenter only as lying along a ray passing through the station at azimuth $\zeta \pm \delta$, where $\delta \approx 7°$. If several stations report such measurements, then, the location is given by the intersection of several sectors of width ~ 14°, and at a distance of $\Delta = 20°$

these sectors are about 550 km wide. The intersection of these sectors is therefore a large area (unless the station-to-epicenter separations Δ are in fact much smaller than 20°), and the location is poorly defined. The addition of back azimuth measurements to a large set of arrival time measurements will thus be expected to offer little improvement in the location beyond that which would be found using the arrival times alone. As we have pointed out earlier in this report, however, this scant improvement should not be taken as demonstrating the worthlessness of back azimuth data for location, since it may still be valuable as an analyst's tool and as a supplement to a sparse data set of arrival times. In the remaining trials to be discussed, we shall investigate these possibilities.

Although we understand why the calculated epicenters for SALMON were so nearly the same for the trials with and without the back azimuth measurements, the underestimation of the confidence region when the back azimuth is included requires further explanation. Since the location changes so little when the back azimuth data are added, it would seem that the error ellipses would also undergo only little change. Figures 3a and 3b, however, show that the F-statistic error ellipse contracts by a factor of about 1.65 in the linear dimensions when the back azimuth data are added to the arrival time measurements. In order to explain this apparent inconsistency, we shall now digress and examine the computation of the confidence regions which are placed around the calculated epicenters.

In order to compute the confidence region around the location, we write the horizontal approximation to the travel-time curve (1) or to the spherical triangle solution (2) as

$$\vec{\delta}_t = \underset{\sim}{B} (\vec{x} - \vec{x}_o) + \vec{\varepsilon} \tag{11}$$

where $\vec{x}$ is the true location in four dimensions and where $\vec{\delta}_t$ is the vector of travel-time residuals computed using the initial guess $\vec{x}_o$ as the location. We now assign weights to each equation in the system (11) by assuming that the error vector $\vec{\varepsilon}$ is a Gaussian zero mean random vector with a covariance matrix given by

$$\text{cov } (\vec{\varepsilon}) = E(\vec{\varepsilon}\vec{\varepsilon}') = \sigma^2 \underset{\sim}{\Sigma}, \tag{12}$$

where $\underset{\sim}{\Sigma}$ is the matrix whose inverse is the matrix of weights

$$\underset{\sim}{\Sigma}^{-1} = \begin{bmatrix} 1/\sigma_1^2 & 0 & \cdots & 0 \\ 0 & 1/\sigma_2^2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 1/\sigma_N^2 \end{bmatrix} . \qquad (13)$$

We have already discussed some of the problems which are presented by the choice of weights for each residual, which we have taken to be $w_i = 1/\sigma_i^2$. Specifically, we noted that in order to make the location less sensitive to large residuals corresponding to outlying data points, the weights w_i should be adjusted at the end of each iteration; we found, however, that this scheme did not work well (whether arrival time and/or back azimuth data were used), and so we adopted the *ad hoc* approach which was described previously. In any event, some weighting function was chosen so that certain observations influenced the location more strongly than did others, and, in accordance with our previous discussion of how location is better defined by using arrival time data rather than back azimuth data, it is justifiable that the back azimuth measurements should be assigned lower weight by the large value which is assumed for the standard deviation, $\sigma = 7°$. Further justification for this assignation of weights will be given later.

The system of N equations (1) has a least-squares solution $\hat{\vec{x}}$ given by

$$\hat{\vec{x}} = \vec{x}_o + (\underset{\sim}{B}^T \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})^{-1} \underset{\sim}{B} \underset{\sim}{\Sigma}^{-1} \vec{\delta} \qquad (14)$$

and a confidence interval may be placed around this estimate of the hypocenter by using an "error ellipsoid" based on either the chi-squared statistic (Evernden, 1969) or the F-statistic (Flinn, 1965, 1969). In the former case, it is assumed that the variance scale factor σ in equation (12) is a known quantity, and it follows that the confidence region consists of all points $\vec{x}$ for which

$$(\vec{x} - \hat{\vec{x}})' \ (\underset{\sim}{B}^T \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B}) \ (\vec{x} - \hat{\vec{x}}) \leq \sigma^2 \ \chi^2_{4:\alpha}, \qquad (15)$$

where $\chi^2_{4:\alpha}$ denotes the level α critical value for a chi-squared random variable with four degrees of freedom. If, on the other hand, σ^2 is believed to be unknown, it may be estimated *a posteriori* by

$$\hat{\sigma}^2 = \frac{1}{N-4} [\vec{\delta} - \underset{\sim}{B} (\hat{\vec{x}} - \vec{x}_o)]' \underset{\sim}{\Sigma}^{-1} [\vec{\delta} - \underset{\sim}{B} (\hat{\vec{x}} - \vec{x}_o)] \tag{16}$$

where we shall use the final iterative value of $\hat{\sigma}^2$ which results when the residuals $\vec{\delta}_t$ are computed using the least-squares location $\hat{\vec{x}}$ as the initial guess $\vec{x}_o$. It then follows that the confidence region is given by all points $\vec{x}$ for which

$$(\vec{x} - \hat{\vec{x}})' (\underset{\sim}{B}^T \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B}) (\vec{x} - \hat{\vec{x}}) \leq 4\hat{\sigma}^2 F_{4,N-4:\alpha}. \tag{17}$$

We see that, as previously asserted, the matrix of weights causes the residuals to be dimensionless, so there is no theoretical distinction between measurements of arrival time and of back azimuth. In Figures 3a and 3b we are concerned with the two-dimensional error ellipses surrounding the calculated epicenter rather than the four-dimensional error ellipsoids surrounding the calculated hypocenter, so we project from four dimensions to two by introducing the matrix $\underset{\sim}{A}$ and the vector $\vec{U}$, where

$$\underset{\sim}{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{18}$$

$$\vec{U} = \underset{\sim}{A} \vec{x}. \tag{19}$$

The chi-squared statistic error ellipse is then given by

$$(\vec{U} - \underset{\sim}{A} \hat{\vec{x}})' [\underset{\sim}{A} (\underset{\sim}{B}^T \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})^{-1} \underset{\sim}{A}^T]^{-1} (\vec{U} - \underset{\sim}{A} \hat{\vec{x}}) \leq \sigma^2 \chi^2_{2:\alpha}, \tag{20}$$

and the F-statistic error ellipse is given by

$$(\vec{U} - \underset{\sim}{A} \hat{\vec{x}})' [\underset{\sim}{A} (\underset{\sim}{B}^T \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})^{-1} \underset{\sim}{A}^T]^{-1} (\vec{U} - \underset{\sim}{A} \hat{\vec{x}}) \leq 2\hat{\sigma}^2 F_{2,N-4:\alpha}. \tag{21}$$

Finally, we note that the "standard error" in the north and the east coordinates (cf. Table IV) is given by the square roots of the first and second diagonal elements of $\sigma^2 \cdot (\underset{\sim}{B}^T \underset{\sim}{\Sigma}^{-1} \underset{\sim}{B})^{-1}$.

Comparing Figure 3a to Figure 3b (and compensating for the change in scales) we see that the chi-squared error ellipses are almost identical for the two cases. This is to be expected, since the right-hand side of equation (20) is the same in both instances and since, as we shall now show, the left-hand side differs only slightly in the two cases. We note that the arrival time measurements are the same in both cases, so the matrix $\underset{\sim}{B}$ for Trial 2 contains the matrix $\underset{\sim}{B}$ for Trial 4 as a submatrix, and we also note that the remaining submatrix of $\underset{\sim}{B}$ for Trial 2, namely the terms involving the back azimuth data, are given small weight by the matrix $\underset{\sim}{\Sigma}^{-1}$. The left-hand side of equation (20) is thus almost the same in both cases, and this is borne out by the similarity of the standard errors for Trials 2 and 4 which are given in Table IV. Since the left-hand side of equation (20) is nearly the same in both cases and since the right-hand side is exactly the same, it follows that the two chi-squared error ellipses are nearly congruent. In both cases they are much too small, since they fail to include (at the 95% confidence level) the true epicenter. This failure indicates that the factor σ^2 on the right-hand side of equation (20) is too small. Since for the sake of convenience we want σ^2 to be unity, the underestimation of the confidence region by the chi-squared error ellipse shows that we should assume a larger *a priori* variance for use in equation (13).

As we have mentioned, in contrast to the near congruence of the chi-squared statistic error ellipses, the F-statistic ellipse in Figure 3a is not the same as is that in Figure 3b but is instead about 1.65 times smaller in its linear dimensions. Since the left-hand side of equation (21) is, for the reasons mentioned above, nearly the same in both cases, the value of the right-hand side must be about $(1.65)^2 = 2.72$ times greater for Trial 4 than for Trial 2. Since the 95% confidence level F-statistics are given by $F_{2,18} = 3.57$ for Trial 2 and $F_{2,8} = 4.46$ for Trial 4 [note that the second number of degrees of freedom in the F-statistic is N-3, rather than N-4, because the depth is restrained; note also that for Trial 2, N = 21 rather than 23 since two back azimuth measurements were deleted by the algorithm because the residuals exceeded 15° (cf. Table IV)], we see that $\hat{\sigma}^2$ is 2.18 times greater for Trial 2 than for Trial 4. How this difference in $\hat{\sigma}^2$ comes about can be seen in the definition of $\hat{\sigma}^2$, equation (16). Since the back azimuth residuals in Trial 2 are assigned low weight, they add little to the sum of the squares of the travel-time residuals, which are the same for Trial 2 and Trial 4. The sums of the squares of the total residuals are thus nearly the same in both cases, and the

main difference in the values of $\hat{\sigma}^2$ then lies in the number of degrees of freedom N-4 which appears as the denominator in the right-hand side of equation (16). We see that the discrepancy lies in the fact that those observations to which low weights were assigned contribute little to the sum of the squares of the residuals, but they contribute to the total number N equally with the higher-weighted observations. Considering the limiting case of adding to a few heavily weighted observations many observations to which negligible weights were assigned, we observe that the left-hand side of equation (21) would be unchanged by the additional observations, $\hat{\sigma}^2$ would decrease as $(N-4)^{-1}$ as the total number of observations N increases, and the resulting decrease in the right-hand side of the equation would be enhanced by a further decrease in the F-statistic. The net effect would be that the confidence region would continue to shrink as the number of observations increases, even though the additional observations were given negligible weight. We should point out that this difficulty does not arise merely when azimuth data are added to arrival time data but that it occurs whenever data of unequal weights are combined, as would be the case for event location based on P_n, P_g, and teleseismic P travel times. We have already stated that we believe that the location process could be improved by assigning weights more rigorously and by adjusting the algorithm which changes those weights iteratively; we herein suggest that further attention ought to be paid to the manner in which the weighting scheme affects the F-statistic error ellipse. A suggestion is made in the Conclusion and Recommendations section. Finally, we point out that if the proper values of the weights $w_i = 1/\sigma_i^2$ are used in equation (13), then the value of σ^2 may be assumed *a priori* to be unity, and the chi-squared statistic error ellipse (which does not suffer from the previously described difficulties because it is independent of N) rather than the F-statistic error ellipse may be taken as the best estimate of the confidence region.

Returning now to the twenty-four location trials, we examine the results of a trial using only back azimuth data and no arrival time data (Trial 5 in Table I). Table IV shows that the absolute error resulting from the use of this data set is in fact larger than that which results from using arrival times and back azimuths (Trial 4), justifying our assignation of heavier weight to the arrival times. We note that since no arrival time measurements were used in Trial 5, not only the depth but also the origin time were restrained during this trial. One arrival time measurement was then added to the data set (Trial 6) in order to increase the number of free variables back to three;

the computed epicenter was unchanged by this addition. On account of the failure of Trials 1 and 3 to result in convergence, Trial 6 and all further trials using arrival times were run with the depth restrained to the surface.

We now compare the results of using data measured at three "distant" (from the true SALMON site) stations with the results of using data measured at three "near" stations. We are interested in the use of networks consisting of only three stations because at regional distances from events in many parts of the world only sparse networks would be available. We wish to see how close to the event these stations should be in order to determine the location accurately. We expect the location to be more accurate when the stations are nearby, since the error associated with the use of a single back azimuth measurement increases approximately linearly with increasing separation Δ. In Trial 7 only the arrival times at the three "distant" stations KN-UT, HN-ME, and EK-NV were used; the resulting location error was 40.5 km. If these three arrival times could be supplemented by back azimuth measurements at all the stations in the entire network (Trial 8), as might happen during a first attempt at location before all the P-arrivals are associated with the proper event, this error is reduced to 32.0 km. In contrast to the situation with the "distant" stations, however, it turned out that if only arrival times from the three "near" stations EU-AL, JE-LA, and BL-WV were used (Trial 9), the iterative process failed to converge. This failure was repeated even when the back azimuths from the entire network were added (Trial 10). This latter failure is particularly significant since Trial 5 showed that the back azimuth data when used by itself resulted in convergence. We thus see that both the traditional location program involving arrival time data and the new program adding back azimuth data may fail when a network consisting of as few as three stations is used. Trial 11 uses both the arrival time and the back azimuth measurements at only the three "distant" stations; the resulting location error, 202 km, is significantly larger than that which resulted from Trial 7 when only the arrival times were used. This trial thus serves to show that the addition of back azimuth measurements made at large distances to a data set of arrival time measurements may deteriorate, rather than enhance, the location. The amount of this deterioration might of course have been less had some other set of "distant" stations been used, but it seems unlikely that the back azimuth residuals would often be small enough to improve the location over that which would result from using only arrival times. We expect this situation to be different when stations "near" to the event are used instead, but we are unable to verify this hypothesis, since Trial 12, which

used arrival time and back azimuth data at the three "near" stations, resulted in the same failure to converge which marred Trials 9 and 10. The confidence regions for Trials 5-12 (except for those cases which failed to converge) are shown in Figures 3c-g. The highly eccentric error ellipse in Figure 3e is due to the poor geometrical distribution of the "distant" network, for which two of the three stations, KN-UT and EK-NV, were close to each other compared to their distance from SALMON. Only the chi-squared error ellipse is plotted on this figure, since the F-statistic error ellipse would be so large that the chi-squared error ellipse would vanish if they both were plotted to the same scale.

The shape and the dimensions of the error ellipse shown in Figure 3e demonstrate that the geometrical distribution of a sparse network of stations can exert a major influence on the location algorithm. In order to investigate this influence further, we have performed the location for SALMON using stations lying within only a small sector of azimuths from the true epicenter. The stations which comprise this azimuthally limited network are EU-AL, BL-WV, BR-PA, DH-NY, LS-NH, and HN-ME, all of which lie within the range of azimuths $41.0° \leq \zeta_o \leq 43.5°$ from SALMON (cf. Figure 2a). This network is perhaps not too unrealistic a simulation of situations which would sometimes be encountered in practice if only regional data were available for event location. When both the arrival time and the back azimuth measurements from this network are used for location (Trial 13), the absolute error is found to be 89.7 km; if only the arrival times are used (Trial 14), the error is only slightly greater, 90.2 km. Even though the calculated epicenter is nearly the same for these two cases, once again the inclusion of the back azimuth measurements has the effect of reducing the size of the confidence region, as is shown in Figures 3h and 3i. If, as in Trial 15, only the back azimuth measurements are used, however, a gross error occurs in the location. As Table IV shows, for this case three of the six back azimuth measurements are assumed to be incorrect and are deleted, and two of the three remaining measurements are assumed to be reversed by 180°, with the result that the epicenter is calculated to be near Syracuse, New York, 1748 km distant from the true site of SALMON. This extraordinary miscalculation serves to demonstrate the difficulties which can result when back azimuth data are used alone for event location, either for "lonesome L_g" regional events or for finding an initial approximation to the epicenter prior to associating P-wave signal arrivals with the event. Comparing the disastrous results of Trial 15 with the better-constrained result of Trial 5, which also used only back azimuth measurements, we see the importance of having a broad azimuthal

distribution of stations if the location is to be performed without arrival time data. Basically the difficulty is that the program is unable to find the point which best represents the intersection of a group of rays which are at best almost parallel to each other and which in actual practice (cf. Figure 2a) may slightly diverge away from each other rather than converge at the true epicenter. For location using only back azimuth data, it is therefore essential that stations such as EK-NV and KN-UT in Figure 2a and SJ-TX in Figure 2b be included in the network. It should be noted that even if there occurred a gross error in location such as that of Trial 15, the confidence regions which are also computed by the algorithm might still serve to identify the resulting epicenters as being possibly spurious. Figure 3j shows that the chi-squared statistic error ellipse is large, and the F-statistic error ellipse (off scale) is far larger. We may thus find some consolation in the knowledge that, at least for this particular case, even though an analyst using the data base of Trial 15 would calculate a grossly inaccurate value for the epicenter of SALMON, he also would know that he should put little faith in this value.

Our principal motivation for introducing back azimuth data into the location algorithm was to enable location to be carried out at regional distances when not enough observations of P-wave arrivals are available to permit the event to be located using arrival time data alone. In order to simulate such circumstances, we now consider a network consisting of only two stations at which both arrival time and back azimuth measurements were made. Considering first the case in which the two stations are "distant" from the event (Trial 16), we choose HN-ME and EK-NV as our "network"; we assume that this distribution of stations represents perhaps the worst possible case which is likely to be encountered in practice. Table IV shows that the absolute error in location which results from the use of these two back azimuths and two arrival time measurements is 955 km. The error ellipse, shown in Figure 3k, is large in one direction, but it fails to include the true epicenter. Had some other two-station "distant" network been chosen instead, the very large location error might have been reduced substantially. For example, if HN-ME and RK-ON are used, the resulting error is only 51.3 km. We anticipate that the regional location procedure, although of apparently limited value when applied to a two-station "distant" network, will prove to be useful when applied to a situation which is more likely to be encountered in practice, namely that of two stations "near" to the event. We investigate this possibility by considering back azimuth and arrival time measurements made at EU-AL and JE-LA (Trial 17). At the

short distance between SALMON and these stations (Δ = 2.2°) errors of a few degrees in the back azimuth measurements will not lead to large errors in the calculation of the epicenter. We find that the absolute error is 34.1 km for this case. We note that the error ellipse for the two-station network, shown in Figure 3ℓ, is highly eccentric. We conclude that the use of back azimuth measurements at regional distances is a valuable supplement to sparse measurements of arrival time and may in fact make it possible to perform event location using only two stations.

Now that we have examined the effects of a sparse network on performing location using both arrival time and back azimuth data, we wish to examine further the effects of such a network on performing event location using back azimuth data alone, as would be done for locating "lonesome L_g" events or for determining a preliminary epicenter with which arrival times could be associated. In order to test whether this technique will work with a sparse data set, we have attempted to locate SALMON using only the back azimuth measurements made by the "distant" and "near" three-station networks which were used in Trials 7-12. We find that the error which results from using the "distant" network (Trial 18) is 376 km. Comparing this result with that of Trial 16, we see that for this particular choice of distant stations a significantly smaller error resulted from using three back azimuths than did from using two back azimuths and two arrival times; we do not anticipate that this will be true in general. In any case, the error is still quite large, but Figure 3m shows that the true SALMON site does in fact lie within the confidence interval which surrounds the computed epicenter. We expect a considerably smaller error to result when the three stations are "near" to the event (Trial 19), and we find that in fact the error is 80.3 km. It is important to note that this trial resulted in a reasonably accurate location even though Trials 9, 10, and 12, which used the arrival time measurements at these same three stations, failed to converge. We thus see that the back azimuth measurements can be quite valuable for performing event location at regional distances even in certain cases for which the location could, in theory, be carried out simply by using the traditional measurements of arrival times.

The possibility of performing event location using only back azimuth data has been addressed previously by Smart (1978). Smart's method is quite different from the technique which has been developed in this report, being based upon a statistical analysis employing the theory of Bernoulli trials to analyze

geometrical clustering patterns of back azimuth measurements. Because this method has no close analog in the treatment of arrival time measurements, it has not been used heretofore in this report. We now wish to compare this statistical technique with our method for calculating event location using only back azimuth data. We have denoted in Table II four stations which have been included in our data set for SALMON but which were not used in the corresponding analysis by Smart (1978). [Smart used one station which is not in our data set, WF-MV, but that back azimuth measurement would be deleted by our algorithm since the residual is almost 90°.] In order to facilitate a comparison between the two methods, we have deleted these four stations from our network and have then repeated the location using back azimuths and arrival times (Trial 20), only arrival times (Trial 21), and only back azimuths (Trial 22). As Table IV shows, the resulting errors in location were, respectively, 34.4 km, 34.9 km, and 418.3 km. It is consistent with our previous discussions of the assignation of weights to back azimuth and arrival time measurements that the inclusion of back azimuth data does not significantly change the computed epicenter relative to that which is found by using arrival time data alone. It is also consistent with our previous discussion that the F-statistic error ellipse for Trial 20 (Figure 3o) is reduced relative to that of Trial 21 (Figure 3p). The very large location error which results from Trial 22 is similar to that which was found for Trial 15 in which only back azimuth measurements were used and all the stations were located within a small sector of azimuths. For Trial 22 the azimuthal range is less restricted than in Trial 15, but most of the stations are still to the northeast of the SALMON site, with three stations being more nearly due north (cf. Figure 2a). For this same set of back azimuth measurements, Smart (1978) calculated the epicenter to be about 200 km distant from the true SALMON site. We thus see that the Bernoulli trials methods results in an error only half as large as that which results from the method employed in this report. We are not prepared to conclude from this single diminution in error that the Bernoulli trials technique is superior, since the relative size of the errors is probably sensitive to the particular data set which is used and thus might be reversed for events other than SALMON. If the Bernoulli trials method does in fact lead to smaller absolute errors in location, it is probably because the algorithm is not influenced heavily by outlying data points. We have made the suggestion earlier in this report that the weighting scheme in our algorithm should be adjusted to assign less weight

to isolated large residuals, and perhaps making this change would lead to closer agreement between the two methods. In any case, it will be necessary to analyze many more events before we are able to say whether the Bernoulli trials or the least squares method is preferable for use in event location involving only back azimuth measurements. If arrival time measurements are also to be used in the location, then the least squares approach must be used.

The final two tests which were performed on the SALMON data set were intended to examine the stability of the solutions which were found previously. Specifically, a poor initial guess of the epicenter was used to begin the first iteration of the location program so that we could test whether it would still converge to the same epicenter as that which had been found by using a better initial guess. Trial 23 used the full data set of arrival time and back azimuth measurements, and it did in fact converge to the same location as did Trial 2. Trial 24 used only the back azimuth data, and it converged to a different epicenter from that which had been found in Trial 5. The reason for this discrepancy is that when the poor initial guess (absolute error $\approx$ 720 km) was used as a trial epicenter, several back azimuth measurements were assumed to be in error since the residuals exceeded 15°, and they were deleted from the data set. If there were no compensating arrival time measurements, the deletion of these back azimuth measurements caused the program to find a different epicenter (for which the absolute error was larger) than it had done when all the back azimuth measurements were used. We note that if any of the iterative locations found by Trial 24 had been sufficiently close to those which were found by Trial 5, the residuals would have decreased to less than the 15° cutoff threshold and the deleted measurements would once again have been included in the data set, so the same final location would have been found. As it turned out, some of the initially deleted back azimuth measurements were in fact reinstated in the data set, but two values were still deleted during the final iteration (cf. Table IV). We conclude that if the location is to be performed using only back azimuth data, it may be worthwhile to use several different trial epicenters as initial guesses in the iterative scheme.

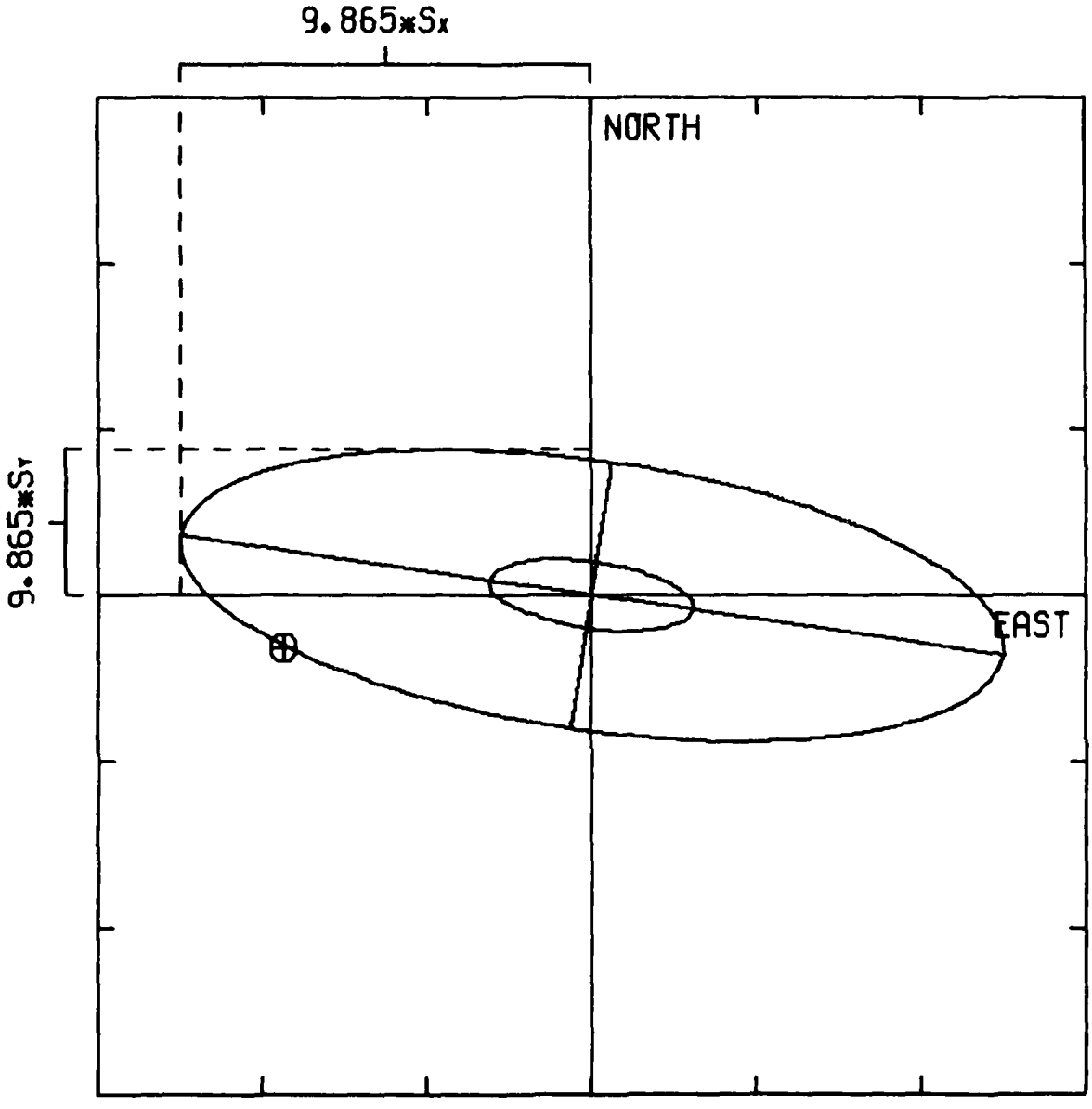
The same twenty-four tests were applied to the data base for GNOME: the results are listed in Table V and are illustrated in Figures 4a-p. As was the case with the SALMON data set, we were unable to perform the location in a depth-free mode, whether or not the back azimuth data were included with the arrival time data (Trials 1 and 3). We have therefore performed all the location tests

TABLE V

Results of the Tests of the Location Algorithm Applied to GNOME

True epicenter and origin time: 32.263°N, 103.865°W, 19:00:00

Trial #	Lat (°N)	Long (°W)	Origin Time	Standard Errors (km) North	Standard Errors (km) East	Absolute Error (km)	No. Azimuths Deleted	No. Azimuths Reversed
1	no convergence							
2	32.321	103.466	19:00:03.3	1.79	5.10	38.12		
3	no convergence							
4	32.322	103.472	19:00:03.2	1.79	5.10	37.59		
5	32.251	101.718	19:00:00.0	48.33	71.37	202.28		
6	32.251	101.718	19:00:14.8	48.32	71.35	202.28		
7	no convergence							
8	31.433	103.569	18:59:58.5	52.35	6.91	96.18		
9	31.881	101.510	19:00:22.5	5.03	20.27	226.32		
10	31.894	101.573	19:00:21.9	4.86	19.60	220.19		
11	no convergence							
12	31.881	101.508	19:00:22.5	4.74	19.03	226.50		
13	33.617	103.093	19:00:12.3	13.05	7.99	166.59	1	
14	33.663	103.069	19:00:12.7	12.80	8.05	172.15		
15	28.376	111.347	19:00:00.0	1344.6	2486.9	838.16		
16	33.530	103.715	19:00:07.2	238.95	59.34	141.19		
17	31.892	101.574	19:00:21.9	12.51	61.19	219.61		
18	33.220	104.022	19:00:00.0	494.68	1369.9	107.12		
19	31.900	101.478	19:00:00.0	53.58	64.67	228.88		
20	32.321	103.466	19:00.03.3	1.79	5.10	38.12		
21	32.322	103.472	19:00:03.2	1.79	5.10	37.59		
22	32.251	101.718	19:00:00.0	48.33	71.37	202.28		
23	32.321	103.466	19:00:03.3	1.79	5.10	38.12		
24	insufficinet data- rejected too many azimuths							6

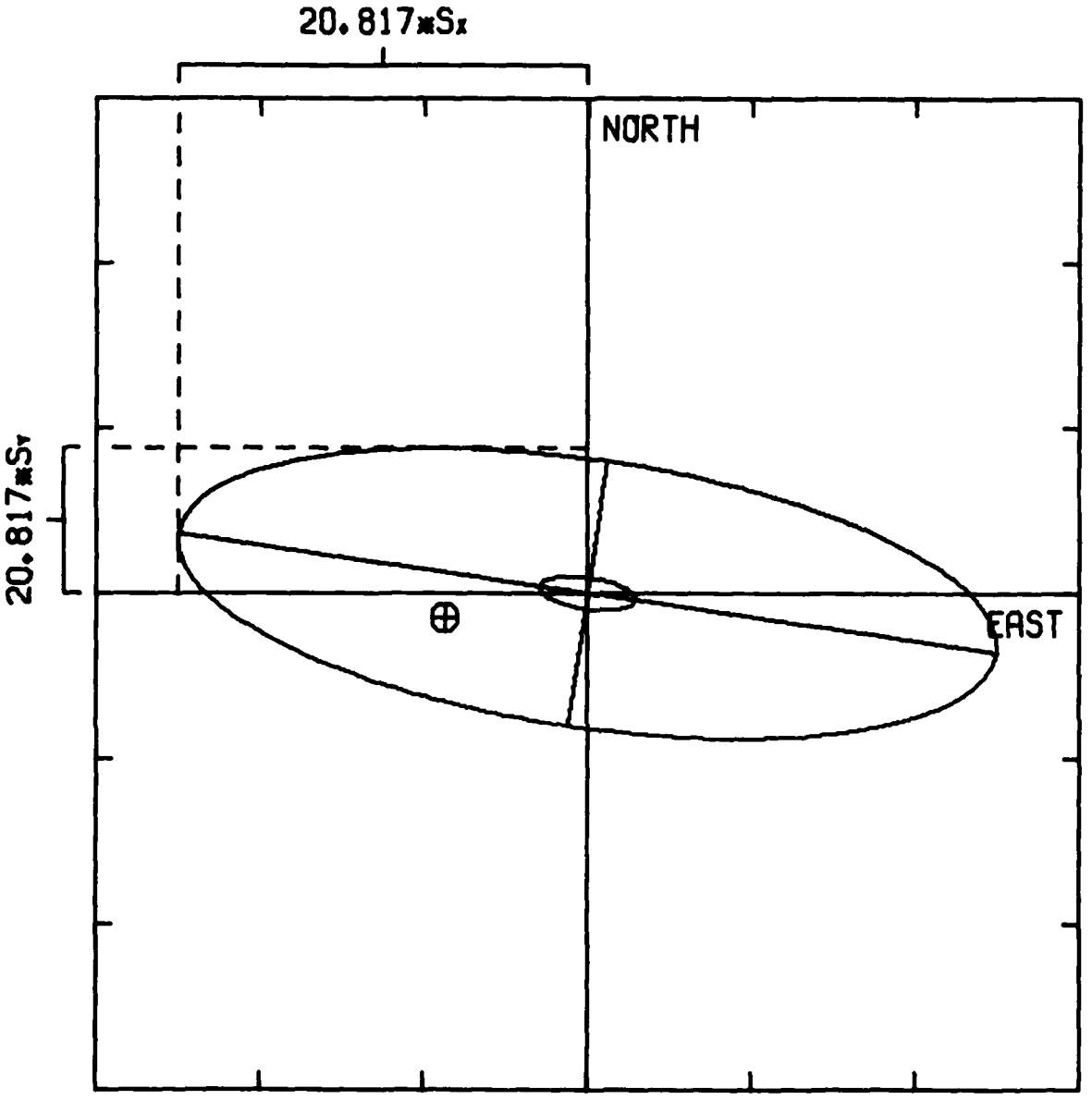


S_x : STANDARD ERROR (EAST)
S_y : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
20.105 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4a Error Ellipses for GNOME, Trial 2

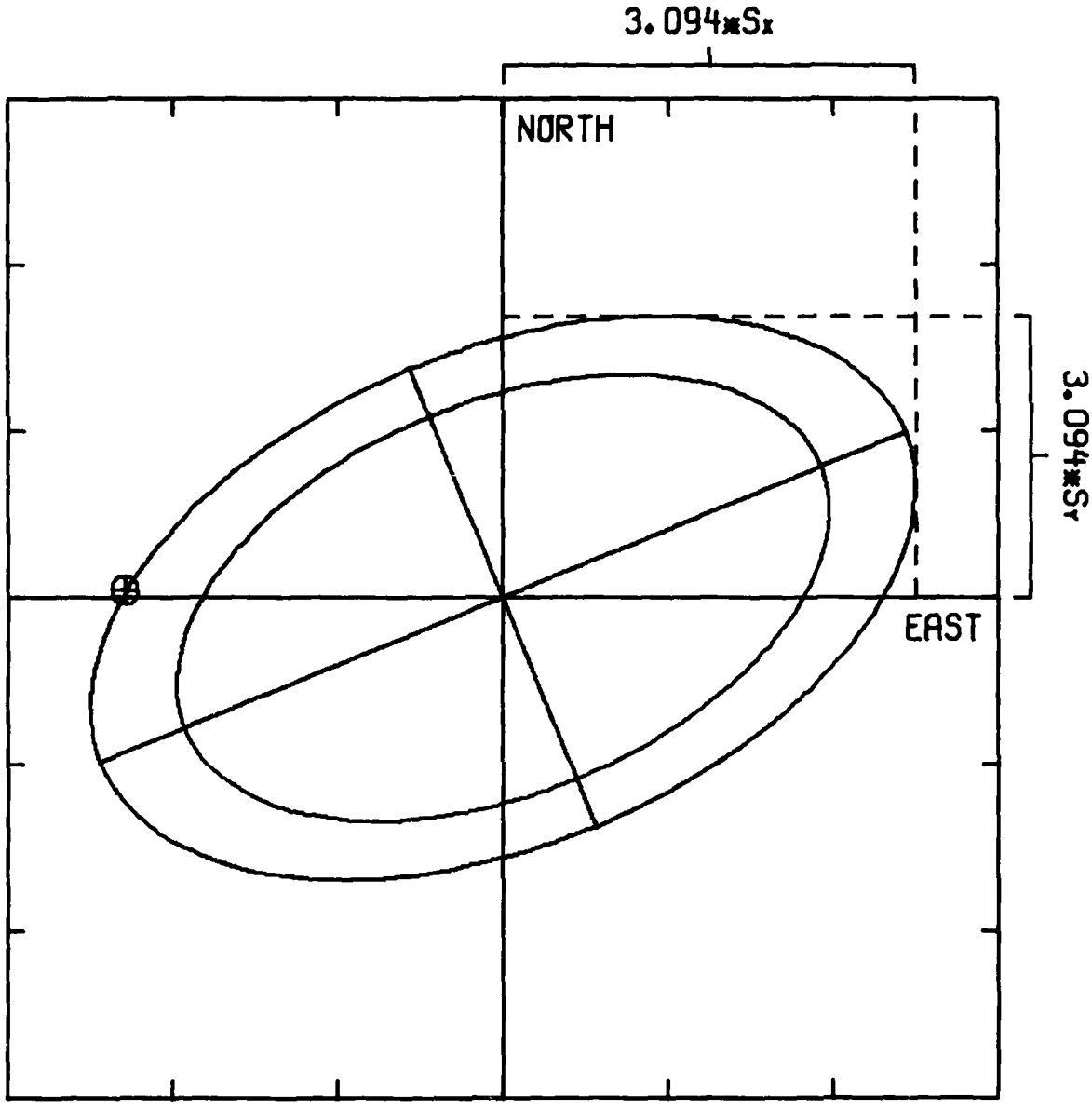


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
42.481 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4b Error Ellipses for GNOME, Trial 4

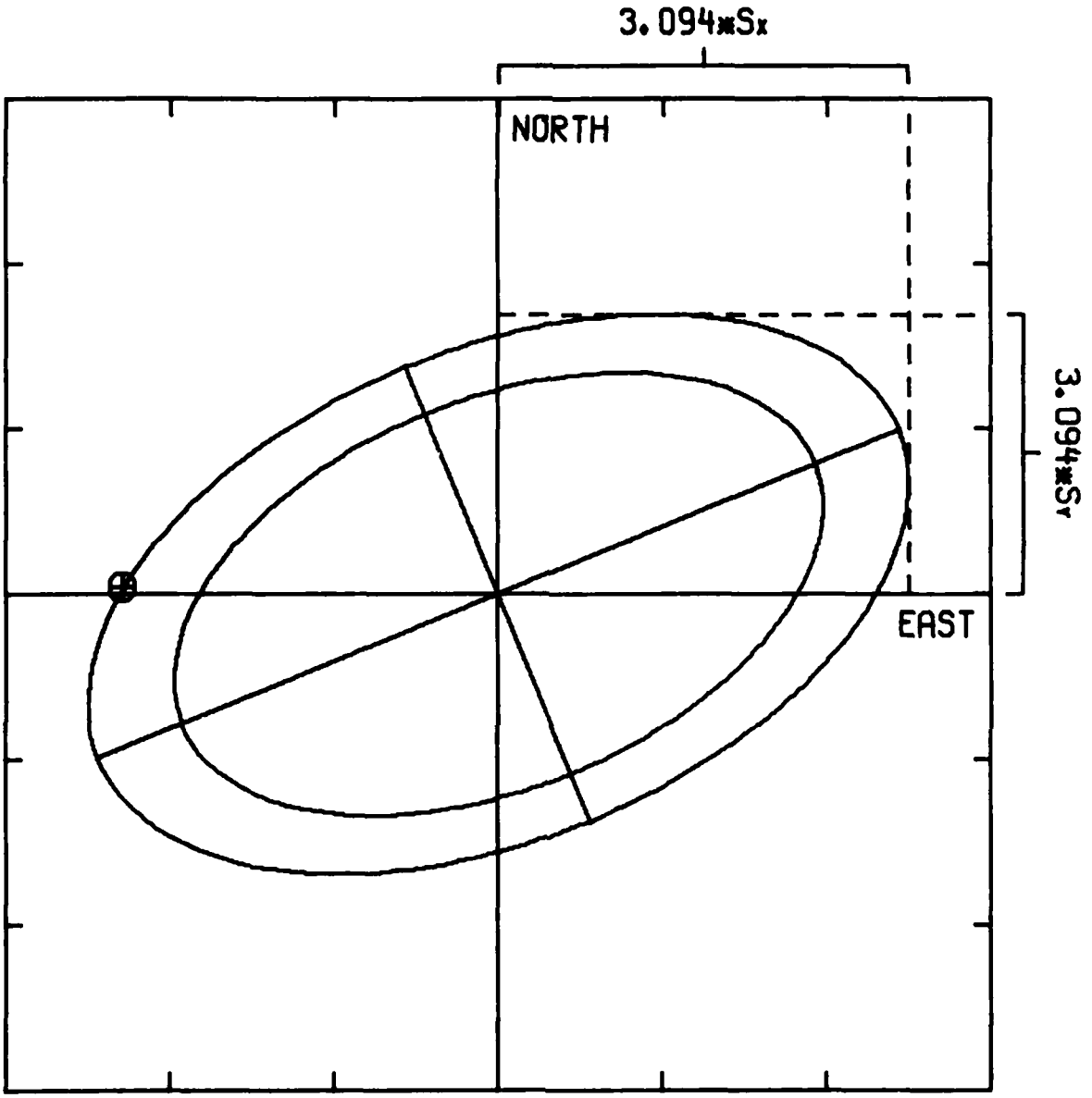


S_x : STANDARD ERROR (EAST)
S_y : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
88.313 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4c Error Ellipses for GNOME, Trial 5

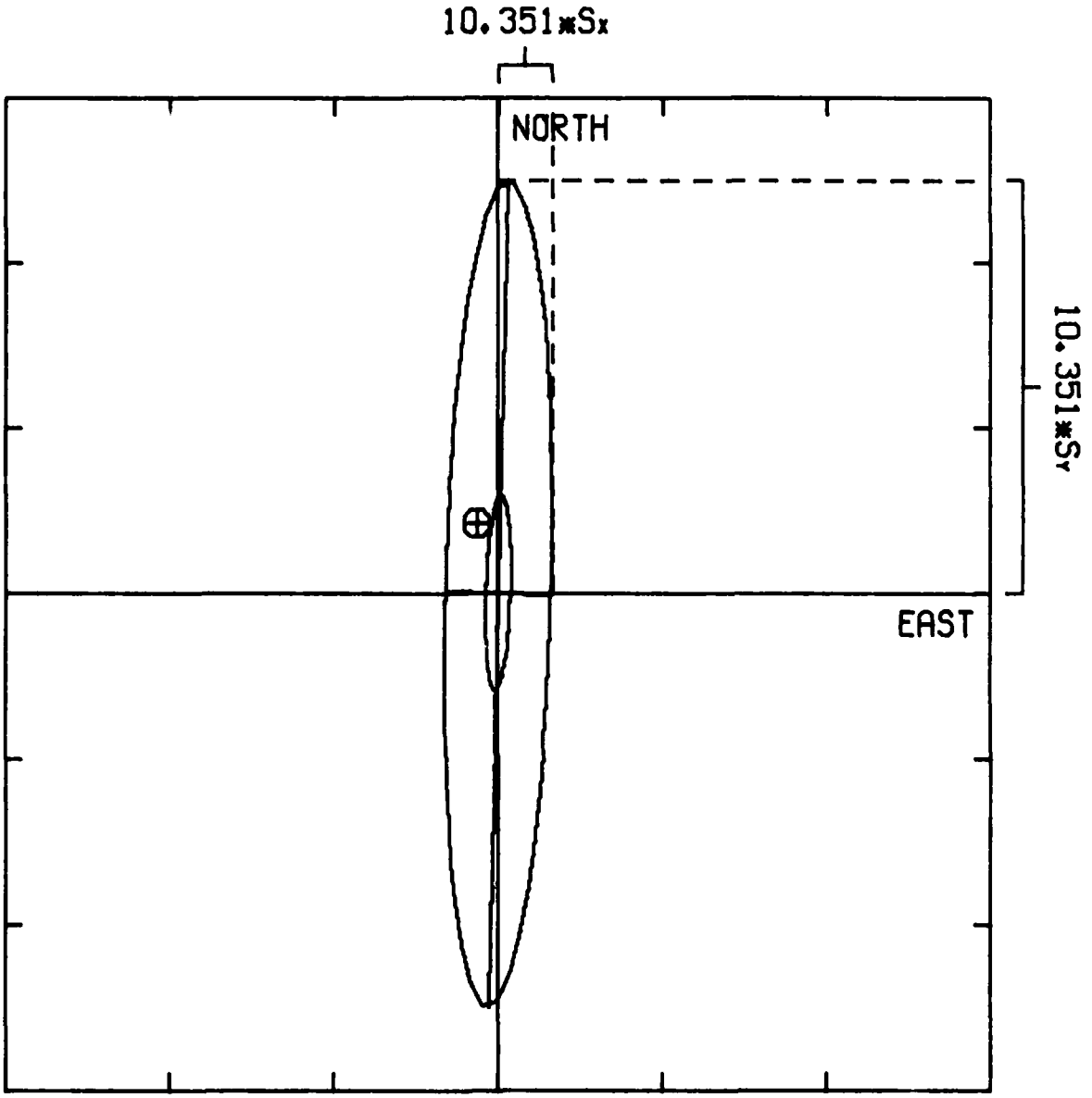


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
88.283 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4d Error Ellipses for GNOME, Trial 6

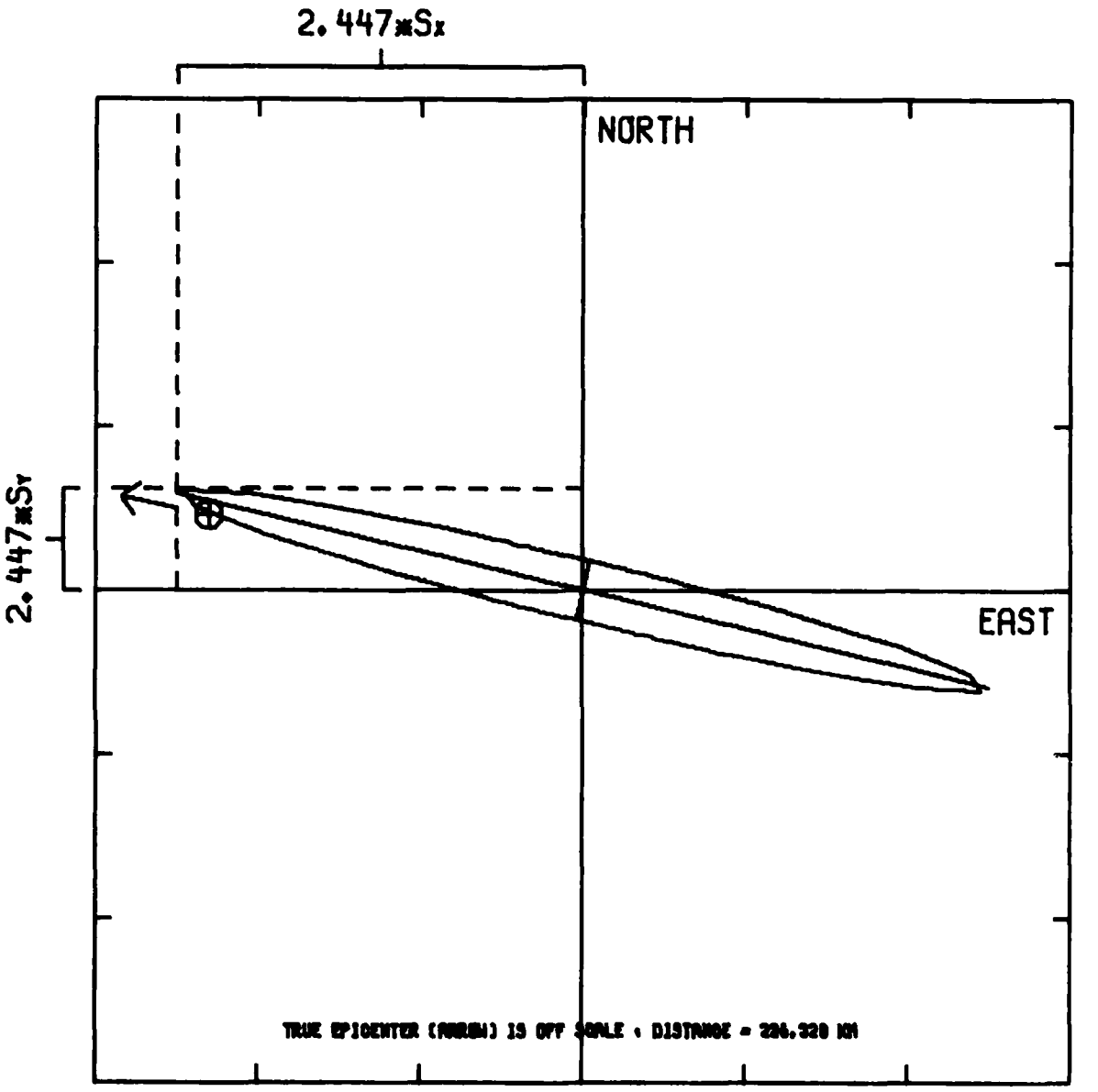


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
216.769 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4e Error Ellipses for GNOME, Trial 8

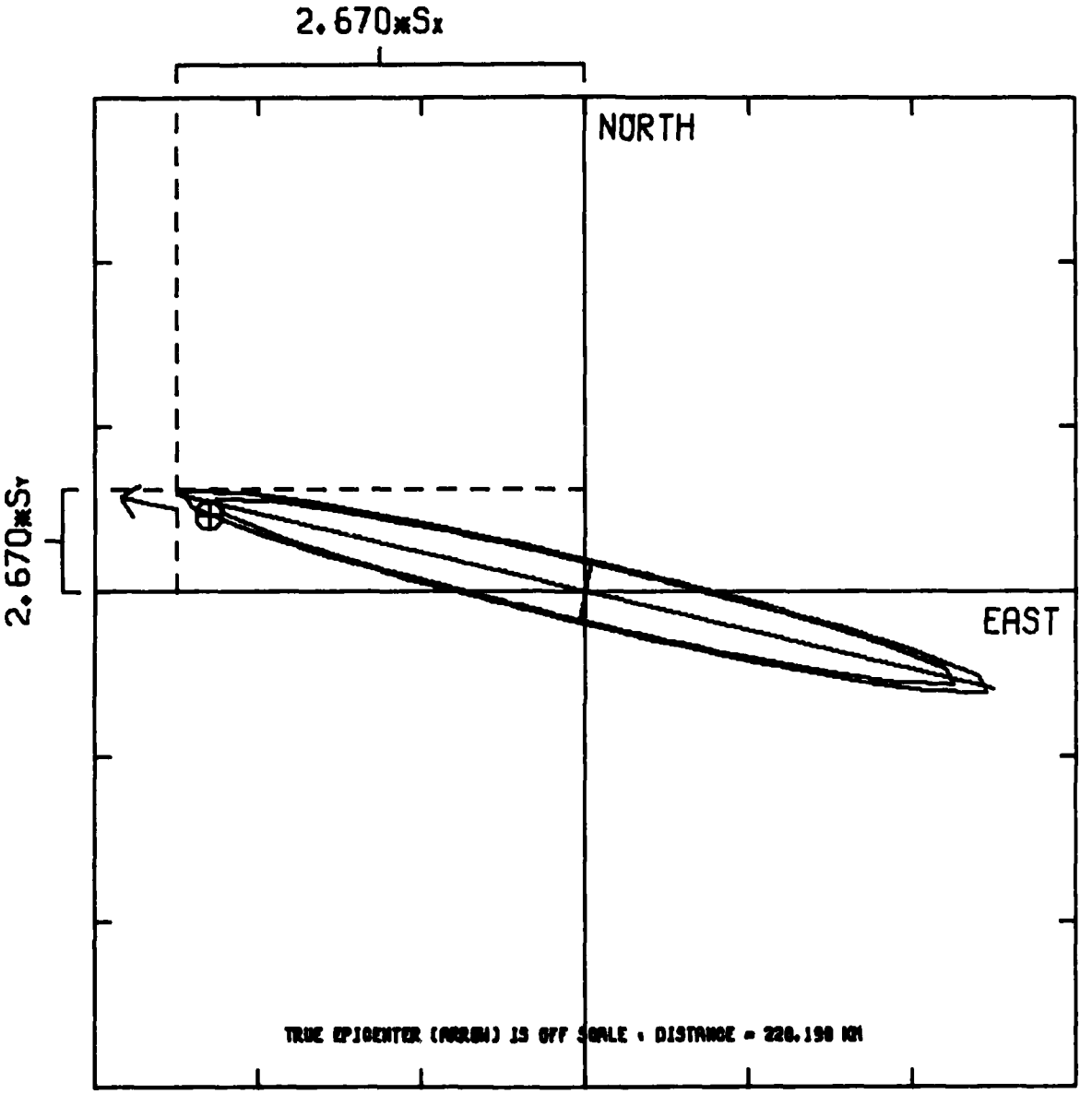


S_x : STANDARD ERROR (EAST)
S_y : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 19.840 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 4f Error Ellipses for GNOME, Trial 9

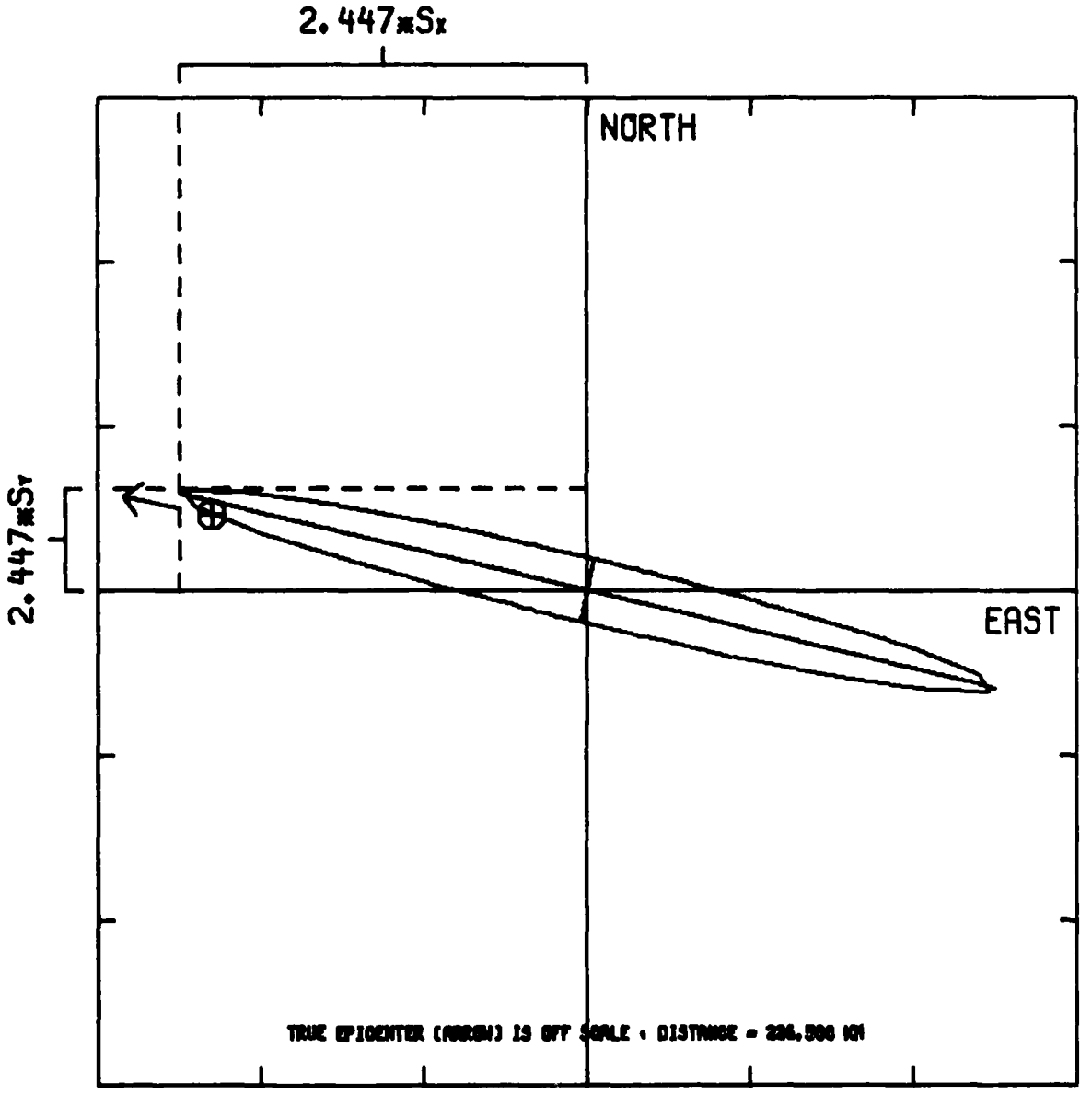


S_x : STANDARD ERROR (EAST)
S_y : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
20.929 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4g Error Ellipses for GNOME, Trial 10

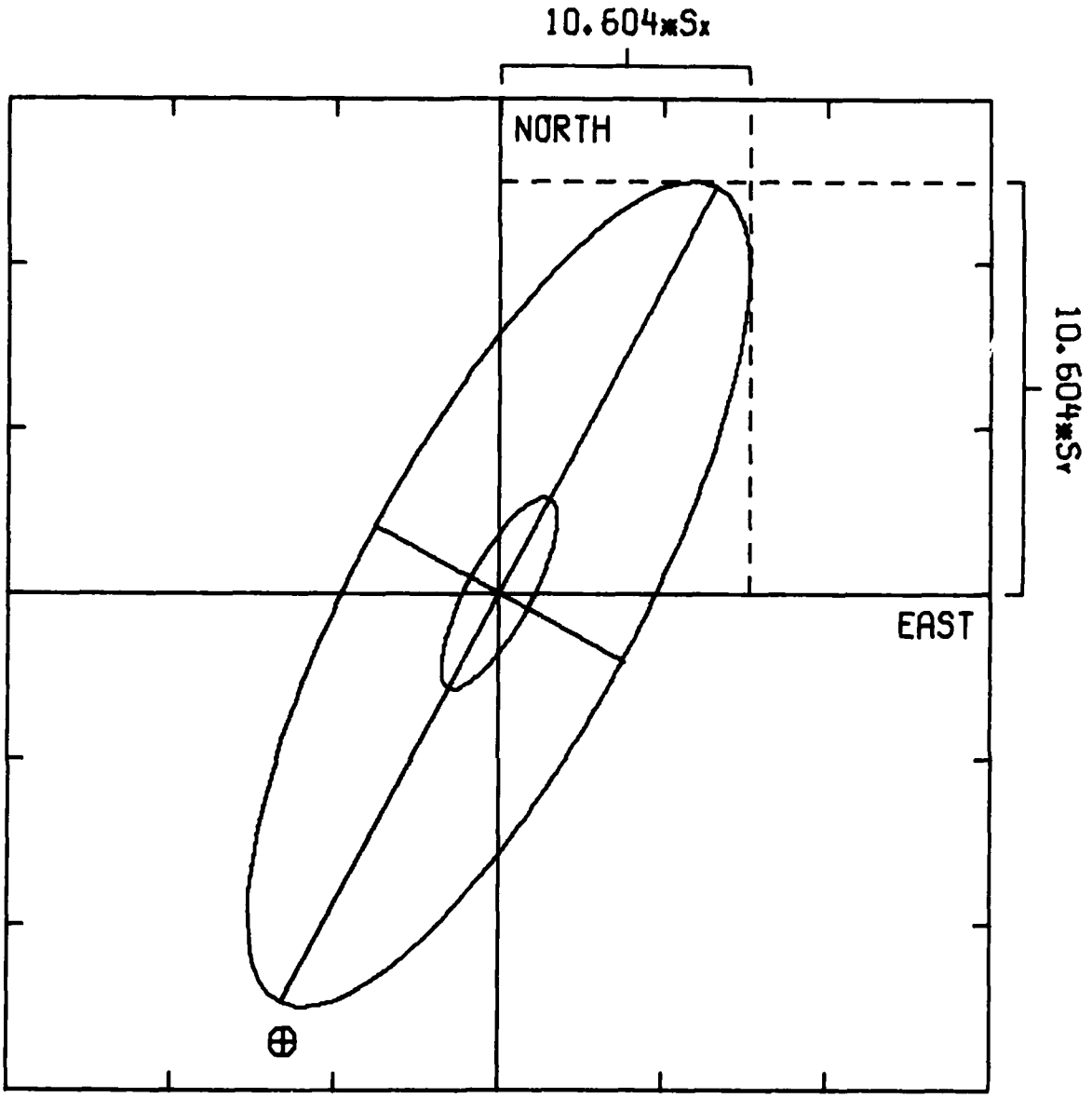


S_x : STANDARD ERROR (EAST)
S_y : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 18.630 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 4h Error Ellipses for GNOME, Trial 12

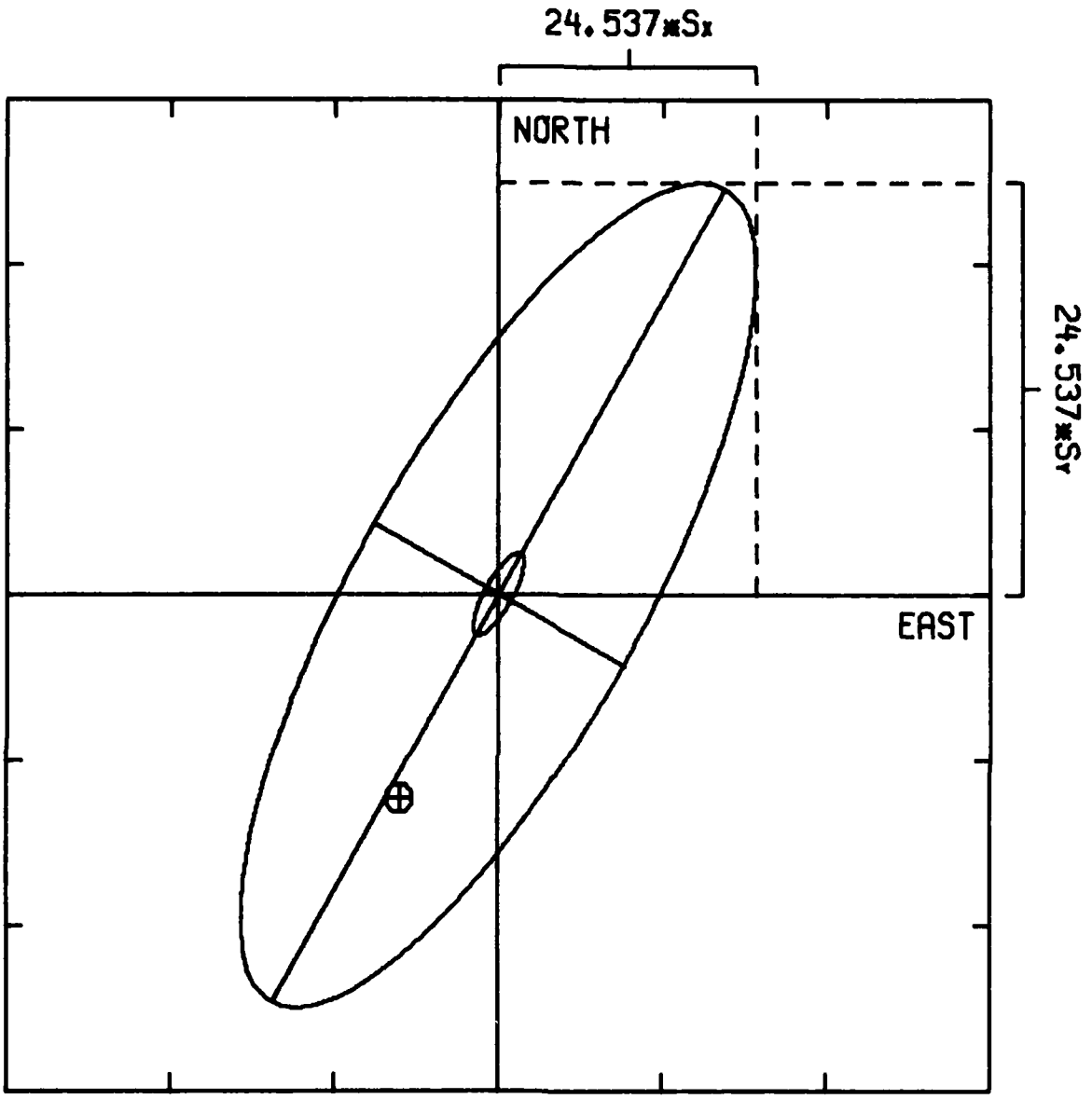


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
55.367 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4i Error Ellipses for GNOME, Trial 13

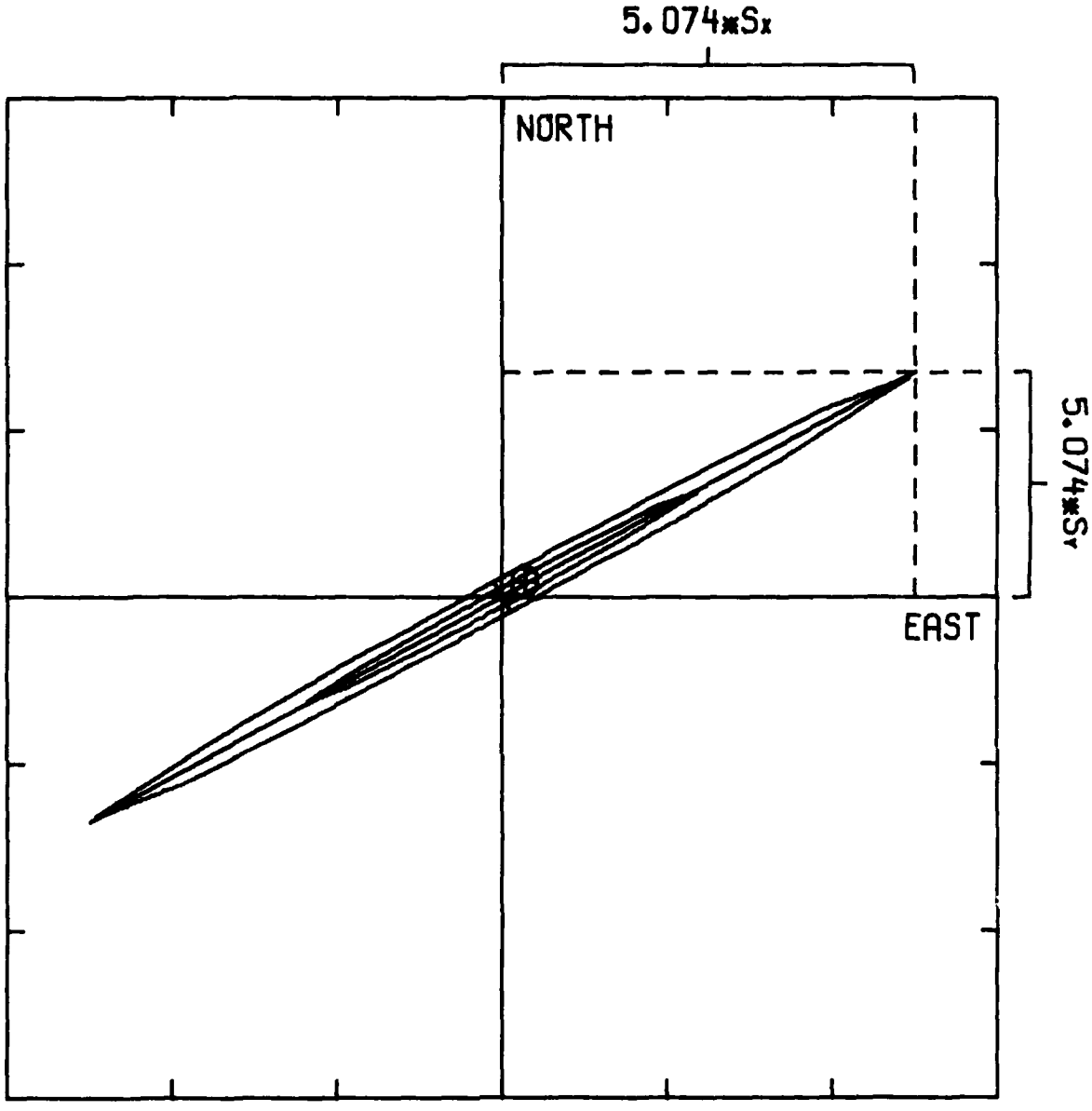


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
125.603 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4j Error Ellipses for GNOME, Trial 14

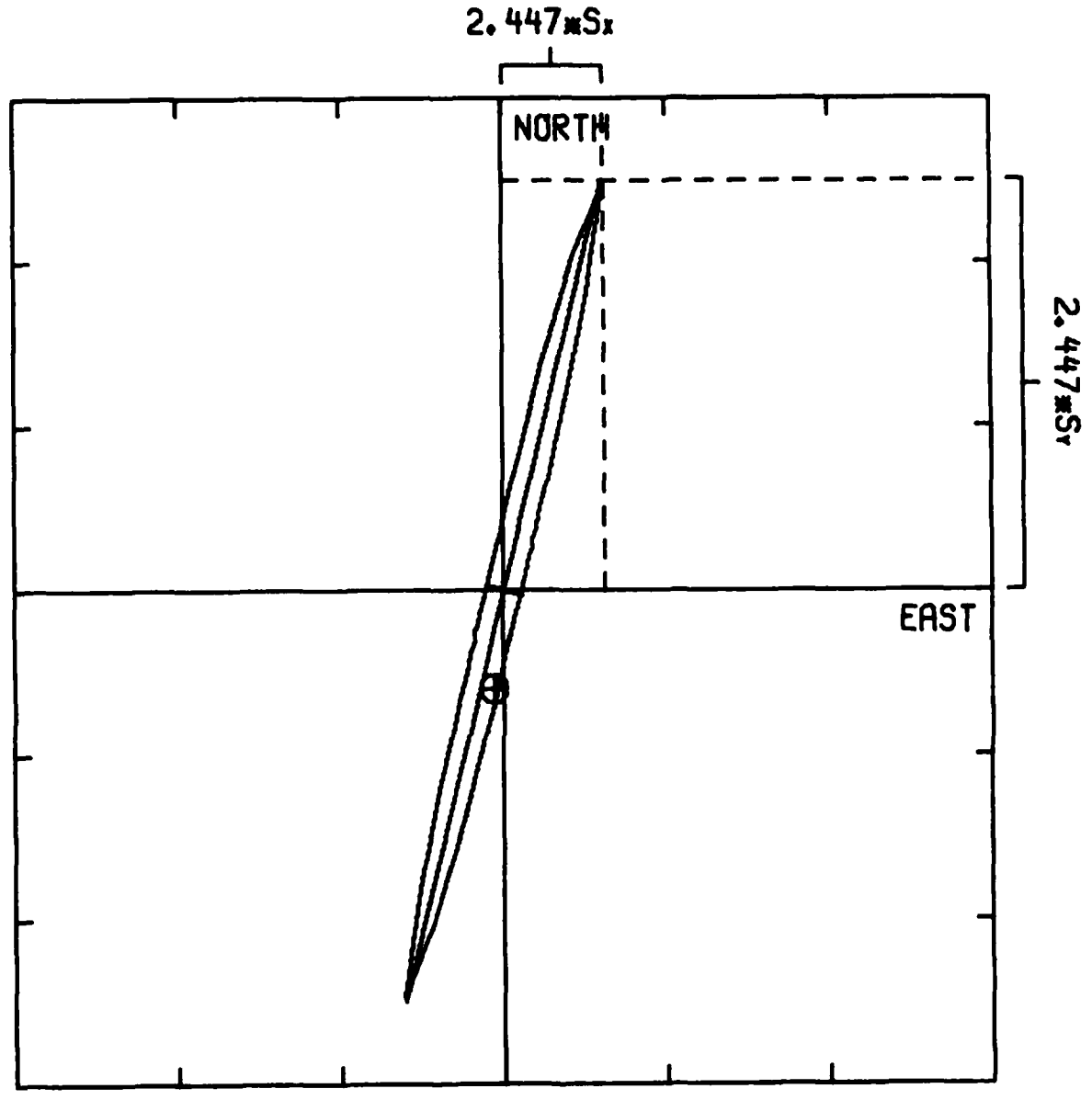


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
5047.851 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4k Error Ellipses for GNOME, Trial 15



Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
233.927 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 41 Error Ellipses for GNOME, Trial 16

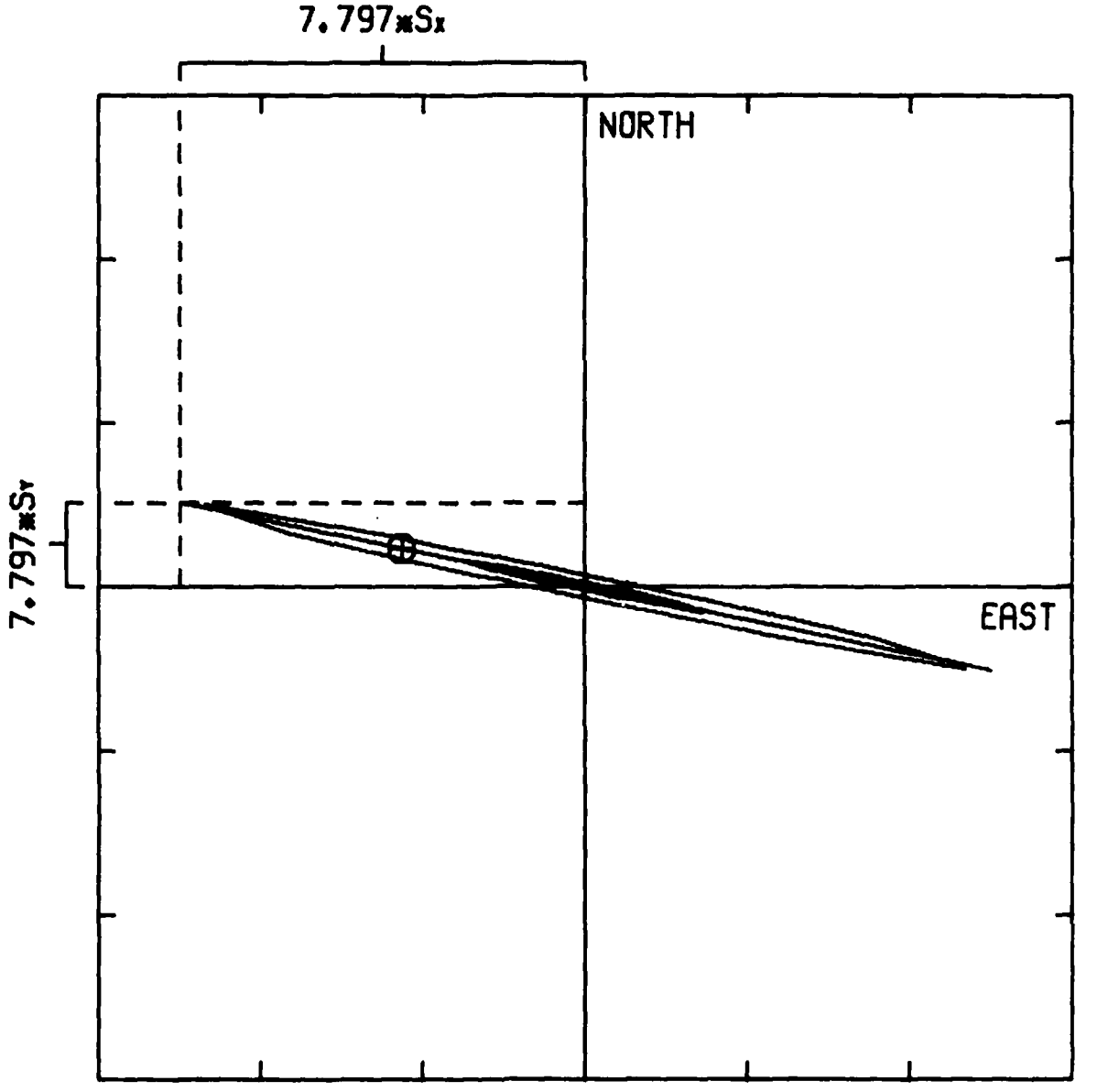
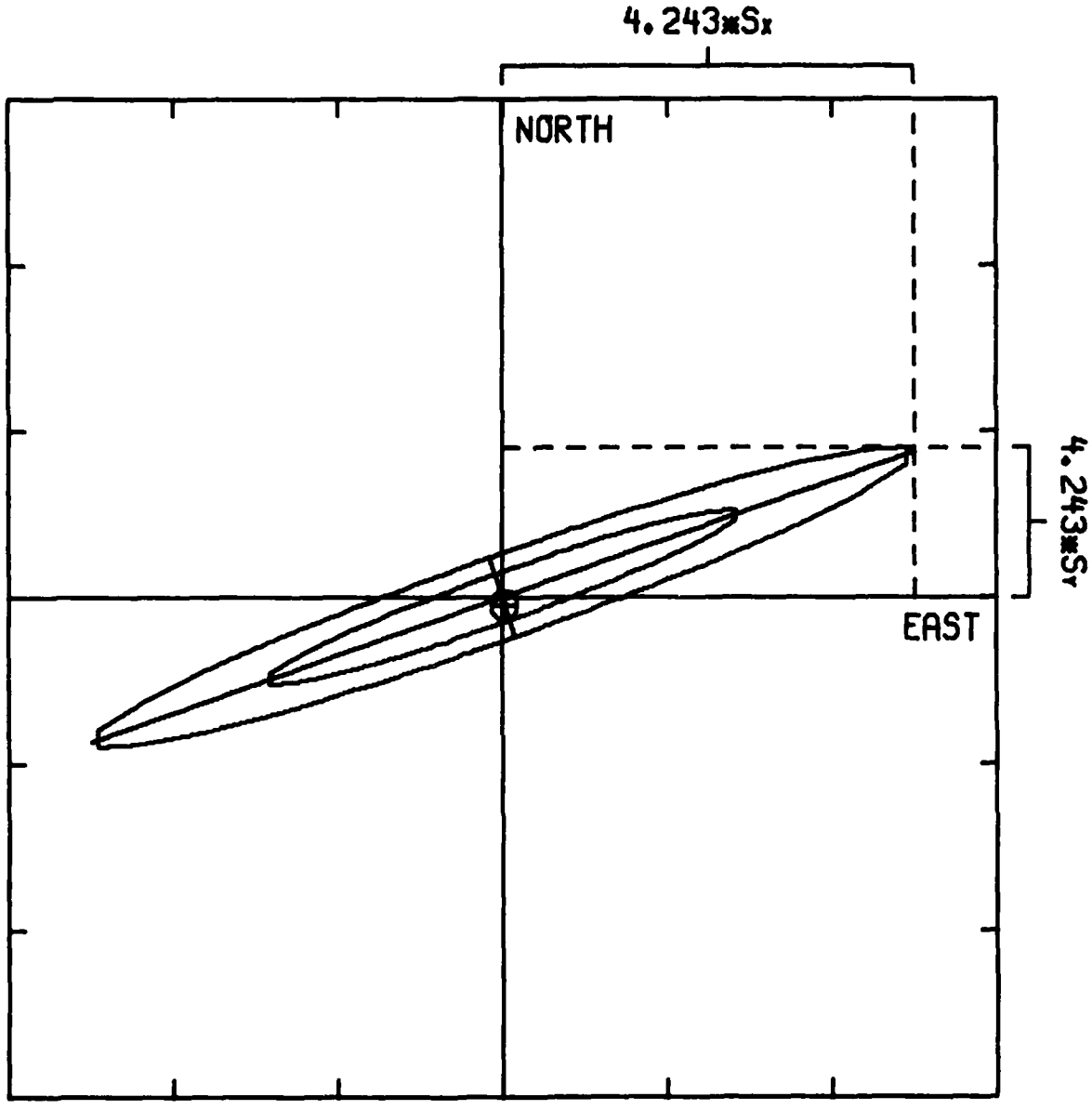


Figure 4m Error Ellipses for GNOME, Trial 17

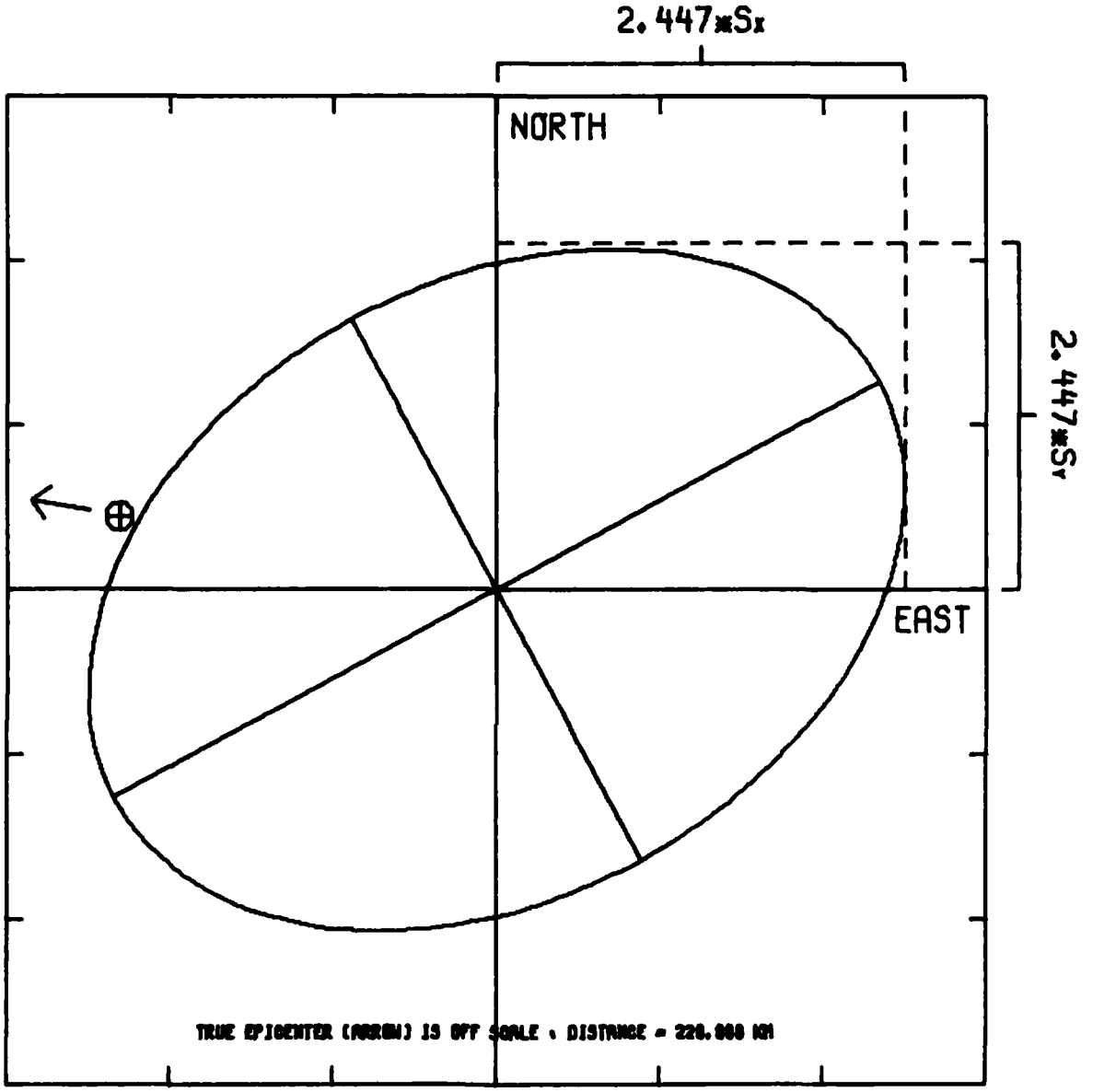


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
2324.797 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4n Error Ellipses for GNOME, Trial 18

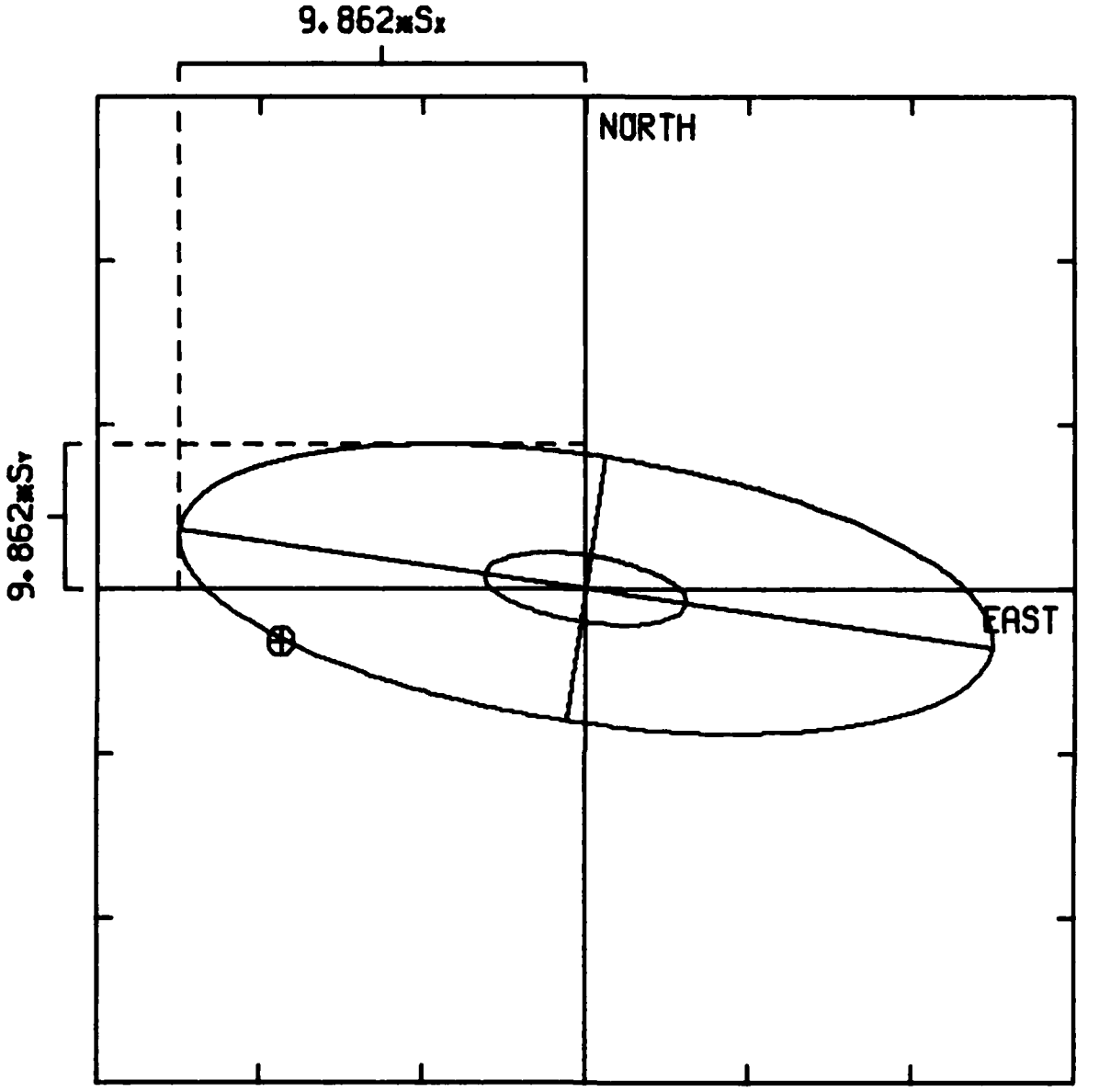


Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
⊢——⊣ 63.307 KM

ELLIPSE IS FOR CHI-SQUARED STATISTIC

Figure 4o Error Ellipses for GNOME, Trial 19



Sx : STANDARD ERROR (EAST)
Sy : STANDARD ERROR (NORTH)

⊕ TRUE EPICENTER
20.099 KM

OUTER ELLIPSE : F-STATISTIC
INNER ELLIPSE : CHI-SQUARED STATISTIC

Figure 4p Error Ellipses for GNOME, Trial 23

with the depth restrained to the surface. When the depth-restrained location is run using both back azimuth and arrival time data (Trial 2), the absolute error is actually greater (cf. Table V) than when only the arrival time data are used (Trial 4). The fact that the location is worsened by the inclusion of the back azimuth measurements is good reason for assigning them low weight. Even though adding the back azimuth worsens the absolute error, it shrinks the F-statistic confidence ellipse, as is shown in Figure 4a and Figure 4b. This contradiction leads us once again to conclude that the back azimuth measurements, while having little influence on the location on account of their weight, have too much influence on the F-statistic error ellipse on account of their number; see the Conclusion and Recommendations section. The figures also show us that we should not trust the chi-squared statistic error ellipse as an alternative, however, since it is too small. As we have seen, this situation may be alleviated by a better choice of the *a priori* standard deviation σ_i for each measurement.

That the inclusion of the back azimuth data increases the absolute error implies that a large error would result if those data were considered alone. Trial 5 shows that this error would in fact be 202 km, over five times larger than the error which results from the use of only arrival time data. Adding one arrival time measurement (made at the most distant station in the network, DH-NY) to the seven back azimuth measurements (Trial 6) failed to diminish this error. The purpose of adding one arrival time measurement to the data set was to increase the number of free variables from two to three by permitting the origin time to be calculated; because the calculated epicenter is so far from the true GNOME site, however, the value which is computed for the origin time is in error by almost 15 sec (cf. Table V). Although the 202 km error for GNOME which resulted from performing the location with only azimuth data substantially exceeds the 73 km error which resulted from the corresponding case for SALMON (cf. Table IV), we note that only seven stations were used for GNOME and twelve stations were used for SALMON. We note that when the number of stations for SALMON was reduced to eight (Trial 22), the error increased to 418 km. That the error for GNOME is about half that value may be attributed to the favorable azimuthal orientation of station SJ-TX relative to the other stations in the network, as is shown in Figure 2b. We shall further examine the effects of the geometrical distribution of stations later.

Comparison of Tables IV and V reveals that formation of three-station subsets of the GNOME network leads to problems in iteration convergence for different combinations of stations than those for which convergence failed for the SALMON network. For GNOME, the iterative process failed to converge when only arrival time measurements from the three "distant" stations CV-TN, BL-WV, and DH-NY were used (Trial 7), but when back azimuth measurements from all seven stations were added (Trial 8), the process converged to an epicenter 96.2 km distant from the true GNOME site. If, as in Trial 11, the back azimuth measurements from only the three "distant" stations are used, however, the location once again fails to converge. This situation is in contrast to that for SALMON, for which it was the subset of three "near" stations which led to the convergence problems. We therefore conclude that the worth for event location of a three-station network depends somewhat less on the distance from the event to the stations then it does on the azimuthal distribution of stations and the accuracy of measurements. It is important to note that it is the measurements of arrival time and not of back azimuth which cause the three-station "distant" network for GNOME to fail to find an epicenter by iteration. That this is so can be seen by examining Trial 18, which uses only the back azimuth measurements for the "distant" stations and which converges to an epicenter 107.1 km distant from the true site. The confidence region for this case is quite large, however (cf. Figure 4n).

Although all tests which were performed using a network of three stations "near" to GNOME (AM-OK, SJ-TX, and MP-AR) resulted in convergence, the location was poor in each case. When only the arrival time data are used from these stations (Trial 9), the resulting error in location is 226.3 km. This error is reduced by only 6.1 km when the back azimuth measurements from all the stations are added in Trial 10. The small size of this error reduction is of course predictable, because the three arrival-time measurements carry more weight than do the seven back azimuth measurements. If arrival time and back azimuth measurements from only the three "near" stations are used (Trial 12), the error is actually increased by 0.2 km over the value which was found in Trial 9 when only the arrival time measurements were used. If only the back azimuth measurements from these stations are used (Trial 19), the error is 228.9 km. Comparing the results of Trials 18 and 19 in Table V and in Figures 4n and 4o, we see that even though the use of the three "near" stations for location using only back azimuth data results in a confidence region which is, understandably, much smaller

than that which results from the use of the three "distant" stations, the absolute error is in fact more than twice as large. This shows that the size of the error which results when only back azimuth data are used cannot be anticipated simply on the basis of the location of the stations, since even close-in measurements of the back azimuth can be substantially in error, on account of scattering of the L_g wavetrain by inhomogeneities distributed along even a short propagation path, or situated strictly locally to the detecting station. Table III indicates that in fact only small residuals in back azimuth result from the measurements made at the distant stations, but that a large residual results from the measurement made at the closest station, AM-OK.

We have mentioned that the poor results obtained from GNOME relative to those obtained for SALMON may be due at least in part to the bias inherent in the azimuthal distribution of the seven stations which were used for locating GNOME. In order to examine the effect of the distribution of stations, we have deleted SJ-TX from the network, leaving six stations which lie within an azimuthal sector 14.2° wide (cf. Table III). When both arrival time and back azimuth data for these stations are used (Trial 13), one back azimuth measurement is deleted, and the resulting error in location is 166.6 km (cf. Table V). If the arrival times are used alone, the error is 172.2 km. If only the back azimuth data are used, a gross error of 838 km results, and an epicenter is found which lies far to the southwest of the true site. The true epicenter does nevertheless lie within the confidence region for this case, which is shown in Figure 4k; the extraordinary dimensions of this region show that the calculated epicenter ought to be regarded as being of little value.

The results of the tests of the two-station location using the two "distant" stations BL-WV and DH-NY (Trail 16) and the two "near" stations AM-OK and SJ-TX (Trial 17) should be compared with those of the corresponding tests for SALMON. The error obtained using the "distant" stations is much smaller, and that obtained using the "near" stations is much larger, for GNOME than for SALMON. The error in Trial 16 is in fact smaller than that in Trial 17 (cf. Table V), so the "distant" measurements are more reliable than the "near" ones for these two pairs of stations in the GNOME network. The unreliability of the epicenter which is computed using the two "near" stations demonstrates the potential dangers inherent in performing event location at regional distances using a sparse data set. It should be pointed out, however, that for both Trials 16 and 17 the true epicenter lies within the F-statistic error ellipses (cf. Figures 4ℓ and 4m).

We do not need to perform Trials 20-22 for GNOME as was done for SALMON, because the network which we used for GNOME is identical to the one which was used by Smart (1978), who obtained a location error of about 160 km using back azimuth data and the Bernoulli trials technique. We therefore proceed to the final two tests, which use a poor initial guess (absolute error $\simeq$ 720 km) to begin the iterative scheme. When the full suite of arrival time and back azimuth measurements are used (Trial 23), the process rapidly converges to the same epicenter as was found for Trial 2. When only the back azimuth data are used (Trial 24), however, the location algorithm fails. What happens in this case is that six of the seven back azimuth measurements are deleted from the data set during the first iteration because the residuals which are calculated using the poorly estimated epicenter exceed the 14° threshold for inclusion, and the process cannot continue. We are thus confronted with a dilemma: our motivation for performing the event location using only back azimuth data is to find a preliminary estimate of the epicenter, but we cannot perform this location without first having a fairly accurate value to use as an initial guess. Unless we try many different values as initial guesses in a suite of location attempts, our only alternative is to drop or at least relax the requirement that the back azimuth residuals be less than 14° at every station. By dropping this requirement, we will in many cases include in the data set misleading measurements which are strongly affected by scattering or which have been associated with the wrong event. It was in order to avoid the influence of these misleading measurements that the 14° cutoff for back azimuth residuals was imposed; we have already argued that it would be preferable to adopt some other weighting scheme (for both back azimuth and arrival time measurements) which would retain outlying measurements in the data base but would assign them low weights. Another possibility is to discard successively the largest residuals, a procedure used in regular location by analysts and by automatic association.

In addition to performing the aforementioned twenty-four tests on the data bases which were compiled for SALMON and for GNOME, we attempted to compile a larger data base of L_g seismograms which could be used for testing the algorithm on additional events. We took as a principal component of this expanded data base the seismograms compiled by Smart and Sproules (personal communication) of thirty-four regional and teleseismic events recorded at the SDCS station RK-ON during the period of March 1976 through February 1978. We therefore undertook to find out how many other SDCS stations (these being the main archival source of three-component short-period digital records of L_g) were operational during

these thirty-four events. We took four stations to be the smallest network which would be considered for any event. It turned out that at least four SDCS stations were operating on the date of twenty-two of the thirty-four events. Eight of these twenty-two events were next eliminated, because either the raw data tapes had been scratched or the data on the tapes were, for various reasons, unusable or a station was not recording at the time of the event. This left fourteen events for which the necessary data records were available from at least four stations. Examination of these data records revealed that only two of the fourteen events were detected at four of the available stations. For both these events, the transverse trace at station NT2NV was inoperative at the time of the event, making the operation of Smart's (1978) three-component processor impossible for that station. We were thus unable to test the location algorithm on any events other than SALMON and GNOME. Further testing must await the compilation of another data base of three-component short-period digital seismograms. Such events certainly exist in the LRSM data base and some may have been digitized already. However, the approach of the end of the contract precluded further data analysis.

CONCLUSIONS AND RECOMMENDATIONS

On the basis of the twenty-four tests of the location algorithm which were performed using the SALMON and GNOME data bases, we have reached several conclusions about the value of incorporating back azimuth measurements into the procedure for locating regional events. We shall now summarize our principal findings.

- In a manner analogous to the handling of back azimuth measurements made by array processing (Julian, 1973), the program LOC can be adapted to handle single-station back azimuth measurements made by Smart's (1978) three-component processor.
- Back azimuth measurements can be used to supplement P-wave arrival time measurements at sparse networks. Reasonably accurate locations were obtained in certain cases using as few as two stations.
- The location algorithm will operate successfully (albeit not always accurately) using back azimuth data alone. It may therefore be used to find the epicenter, but not the depth or origin time, for "lonesome L_g" events as well as to find a trial epicenter with which the proper P-wave arrivals can be associated.
- Two examples of a sparse network (the three "near" stations for SALMON and the three "distant" stations for GNOME) were found for which the iterative location scheme failed to converge when arrival time measurements were used, but which resulted in convergence when back azimuth data were used instead. It may thus be worthwhile to calculate measurements of back azimuth even for networks for which the arrival time measurements should, in theory, be sufficient for location.
- Large absolute errors in location can result, particularly if few arrival time measurements are used, when the network of stations is poorly distributed azimuthally about the event. If only back azimuth measurements are used, large errors can result if the stations are distant from the event. Large errors can also result from measurements made at even short distances if the (true-measured) back azimuth residuals are large.
- On the whole, smaller errors in location result when arrival time measurements are used than when only back azimuth data are used. Adding back azimuth measurements to a large data base of arrival times has little effect on the resulting calculation of the

epicenter. If the influence of the back azimuth data were to be increased by assigning them heavier weight, the resulting absolute error of location would frequently be worsened rather than improved became the precise estimate available from arrival time data would be devalued.

- On account of *a priori* underestimation of the variance of travel time measurements made at regional distances, the chi-squared statistic error ellipse surrounding the calculated epicenter is usually too small. The F-statistic error ellipse may also be too small if the data base of arrival times is supplemented by many azimuth measurements, since even though these low-weighted observations have little influence on the epicenter calculation, they cause the error ellipse to shrink substantially. Qualitatively, this is due to the fact that the assumption of a constant ratio between azimuth and travel-time variance (cf. equations (12) and (13)) causes multiple observations of azimuth to apparently reduce the imprecision in knowledge of arrival time variance. Since arrival time data leads to small confidence ellipses if the variance is known, multiple azimuth observations will lead to a rapid shrinking of the travel-time dominated F-ellipse. The solution to this problem must be to estimate two unrelated variances, σ_θ and σ_t, simultaneously. However, this requires an expansion of the theory.

- A poor first approximation in the iterative location scheme may cause some back azimuth measurements to be judged erroneously as being faulty or as being in error by 180°.

On the basis of this study, the authors make the recommendations listed below for research on improving location at regional distances using back azimuth measurements.

- No tests of the location technique presented in this report have yet been performed successfully in the depth-free mode. More analysis should be performed to measure the effectiveness of the method when the depth coordinate is a free variable. This analysis should use earthquakes at various depths as well as near-surface explosions.

- An examination should be made of the effect of assigning a higher *a priori* estimate of arrival time variance to measurements made at regional distances.

- A more effective method of weighting the data should be employed which would adjust the weights during each iteration so that outlying measurements of the back azimuth would not strongly influence the location, but so that they would also not be assumed to be erroneous and then be completely ignored.

- A large suite of events should be used to carry out a comparison between the method developed in this report for performing event location using only back azimuth data and the Bernoulli trials technique of Smart (1978).

- Further studies should be made of Smart's (1978) three-component processor in order to determine the most reliable method of calculating the back azimuth measurements which are used by the regional location algorithm. In particular, the possibility of using P waves, rather than L_g, should be examined. Recent results have shown that P wave azimuths are reliable and have no 180° ambiguity.

- The effect (on both the calculated epicenter and the confidence region) of including the arrival times of additional seismic phases should be evaluated. Arrival times of P_g and L_g and the arrival time differences P_g-L_g should be tested.

- Although the algorithm has been developed for performing location at regional distances, it should be further tested by adding to the data base teleseismic measurements of P-wave arrival times and P-wave azimuths.

- In order to test further the use of back azimuth measurements in performing event location, it is imperative that a large data archive be constructed which comprises three-component short-period digital seismograms of both P and L_g for a large number of events and for as many stations as possible.

ACKNOWLEDGEMENT

The authors have benefitted in this work from several conversations with Eugene Smart on the topic of the Bernoulli Trials technique of location using back azimuth data.

REFERENCES

Chang, A. C., Rivers, D. W., and J. A. Burnetti (1980). Improved location with regional events, (in preparation), Teledyne Geotech, Alexandria, Virginia.

Chang, A. C., and D. P. J. Racine (1979). Evaluation of location accuracy using P_n and P_g arrivals, SDAC-TR-79-4, Teledyne Geotech, Alexandria, Virginia.

Evernden, J. F. (1969). Precision of epicenters obtained by small numbers of world-wide stations, Bull. Seism. Soc. Am., 59, 1365-1398.

Flinn, E. A. (1965). Confidence regions and error determinations for seismic event location, Rev. Geophys., 3, 157-185.

Flinn, E. A. (1969). Erratum, Rev. Geophys., 7, 664.

Julian, B. R. (1973). Extension of standard event location procedures, Seismic Discrimination SATS, Lincoln Laboratory, M.I.T. (30 June 1973), DDC AD-766559.

Smart, E. (1978). A three-component, single-station, maximum-likelihood surface wave processor, SDAC-TR-77-14, Teledyne Geotech, Alexandria, Virginia.